Draft Comparative Effectiveness Review

Number xx

Cervical Degenerative Disease Treatment: A Systematic Review

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
5600 Fishers Lane
Rockville, MD 20857
www.ahrq.gov

Contract No. To be added to the final report.

Prepared by:
To be added to the final report.

Investigators:
To be added to the final report.

AHRQ Publication No. xx-EHCxxx
<Month Year>

This report is based on research conducted by the XXXXX Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. XXX-20XX-XXXXX). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

None of the investigators have any affiliations or financial involvement that conflicts with the material presented in this report.

The information in this report is intended to help healthcare decision makers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of healthcare services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

A representative from AHRQ served as a Contracting Officer's Representative and reviewed the contract deliverables for adherence to contract requirements and quality. AHRQ did not directly participate in the literature search, determination of study eligibility criteria, data analysis, interpretation of data, or preparation or drafting of this report.

AHRQ appreciates appropriate acknowledgment and citation of its work. Suggested language for acknowledgment: This work was based on an evidence report, [TITLE], by the Evidence-based Practice Center Program at the Agency for Healthcare Research and Quality (AHRQ).

Suggested citation: To be added to final report.

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of systematic reviews to assist public- and private-sector organizations in their efforts to improve the quality of healthcare in the United States. These reviews provide comprehensive, science-based information on common, costly medical conditions, and new healthcare technologies and strategies.

Systematic reviews are the building blocks underlying evidence-based practice; they focus attention on the strength and limits of evidence from research studies about the effectiveness and safety of a clinical intervention. In the context of developing recommendations for practice, systematic reviews can help clarify whether assertions about the value of the intervention are based on strong evidence from clinical studies. For more information about AHRQ EPC systematic reviews, see https://effectivehealthcare.ahrq.gov/about/epc/evidence-synthesis

AHRQ expects that these systematic reviews will be helpful to health plans, providers, purchasers, government programs, and the healthcare system as a whole. Transparency and stakeholder input are essential to the Effective Health Care Program. Please visit the website (www.effectivehealthcare.ahrq.gov) to see draft research questions and reports or to join an e-mail list to learn about new program products and opportunities for input.

If you have comments on this systematic review, they may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 5600 Fishers Lane, Rockville, MD 20857, or by email to epc@ahrq.hhs.gov.

Robert Otto Valdez, Ph.D., M.H.S.A.
Director
Agency for Healthcare Research and Quality

Therese Miller, DrPH
Acting Director
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Craig A. Umscheid, M.D., M.S.
Director
Evidence-based Practice Center Program
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Kim Wittenberg, M.A.
Task Order Officer
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Acknowledgments

The authors gratefully acknowledge the following individuals for their contributions to this project: To be added in the final report.

Key Informants

In designing the study questions, the EPC consulted several Key Informants who represent the end-users of research. The EPC sought the Key Informant input on the priority areas for research and synthesis. Key Informants are not involved in the analysis of the evidence or the writing of the report. Therefore, in the end, study questions, design, methodological approaches, and/or conclusions do not necessarily represent the views of individual Key Informants.

Key Informants must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any conflicts of interest.

The list of Key Informants who provided input to this report follows:
To be added in the final report.

Technical Expert Panel

In designing the study questions and methodology at the outset of this report, the EPC consulted several technical and content experts. Broad expertise and perspectives were sought. Divergent and conflicted opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Therefore, in the end, study questions, design, methodologic approaches, and/or conclusions do not necessarily represent the views of individual technical and content experts.

Technical Experts must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

The list of Technical Experts who provided input to this report follows:
To be added in the final report.

Peer Reviewers

Prior to publication of the final evidence report, EPCs sought input from independent Peer Reviewers without financial conflicts of interest. However, the conclusions and synthesis of the scientific literature presented in this report do not necessarily represent the views of individual reviewers.

Peer Reviewers must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential non-financial conflicts may be retained. The TOO

and the EPC work to balance, manage, or mitigate any potential non-financial conflicts of interest identified.

The list of Peer Reviewers follows:
To be added in the final report.

Cervical Degenerative Disease Treatment: A Systematic Review

Abstract

Objectives. Cervical degenerative disease (CDD) is a common disease that becomes more prevalent with age, with management including surgical and nonoperative treatments to alleviate pain, improve neurologic function, and prevent progression or recurrence. In 2009, the Congress of Neurological Surgeons published guidelines on the management of CDD. This systematic review summarizes the evidence on treatments for CDD in patients with or without cervical radiculopathy or myelopathy, including nonoperative management compared with operative management, which was not part of the 2009 guidelines.

Data sources. We searched Ovid MEDLINE®, EMBASE, and Cochrane CENTRAL from 1980 to July 1, 2022, as well as reference lists and clinical trial registries. Additionally, we reviewed all studies included in the 2009 guidelines for inclusion in this review.

Review methods. Predefined criteria were used to select randomized controlled trials and nonrandomized studies of interventions that addressed the effectiveness and harms of treatments in patients with cervical degenerative disease. Prespecified methods were used to assess individual study quality and strength of evidence for key outcomes. Effects were analyzed using qualitative methods and quantitative synthesis where appropriate.

Results. We included 56 randomized controlled trials and 49 nonrandomized studies. Studies enrolled patients with radiculopathy and/or myelopathy with disease at one or more levels. A variety of surgical approaches and techniques were employed; however, there were few comparative studies that included nonoperative treatments. Most studies were rated moderate risk of bias, while the majority of evidence was rated low or of insufficient strength to draw firm conclusions on comparative benefits and harms.
Arthroplasty versus Anterior Cervical Discectomy and Fusion (ACDF): There were no differences between arthroplasty and ACDF in improved postoperative pain or function across various measures and timepoints (SOE: Moderate at 1-level intervention; SOE: Low at 2-level intervention). However, arthroplasty was associated with a substantially decreased likelihood of reoperation at 24 months with both 1-level surgery (SOE: High) and 2-level surgery (SOE: Low), with similar results at longer followup times.
Anterior versus Posterior approach: There was insufficient evidence to determine comparative benefits for most outcomes between anterior and posterior approaches, with similar reoperation rates across 3 RCTs in patients with radiculopathy and single-level disease (SOE: Low); the likelihood of experiencing any serious adverse event was higher with posterior approaches than with ACDF in patients with 3 or more level disease (SOE: Low; data not pooled due to heterogeneity in reporting specific adverse events across multiple studies).
Standalone cage versus Plate and cage in ACDF: There was no difference in fusion rates between the two fusion techniques at 12 months, with similar fusion rates at 24 and 36 months (SOE: Moderate). There were also no differences between a standalone cage and plate and cage

in postoperative improvement in arm pain, function, quality of life, or adjacent-level ossification following ACDF (SOE: Low). Few reoperations were reported.
Laminoplasty versus laminectomy and fusion. There were little differences between surgical techniques in postoperative neurologic function (SOE: Moderate) or general function (SOE: Low), but the risk of experiencing a complication was lower with laminoplasty (SOE: Low), with no difference in reoperation rates (SOE: Moderate).
Other comparisons. Evidence for other comparisons was limited. No studies meeting inclusion criteria were available to guide management of CDD in asymptomatic patients with radiographic spinal cord compression or to guide management of pseudarthrosis after anterior cervical fusion.

Conclusions. There were few differences in benefits between surgical approaches and techniques compared in included studies for the treatment of cervical degenerative disease. However, there were some differences in the frequency of adverse events for some comparisons. There was substantial evidence that the risk of reoperation is much lower for artificial disc replacement than ACDF. Limited evidence also suggests a lower likelihood of experiencing any serious adverse event with ACDF than PCDF and a lower risk for any complication with laminoplasty compared with laminectomy and fusion. There was limited evidence on the role of nonoperative management instead of surgery or in addition to surgery to treat CDD, and no evidence to determine benefits and harms of a revision anterior arthrodesis or posterior approach in patients with pseudarthrosis after prior anterior cervical fusion.

Contents

Tables

Figures

Appendices

Executive Summary

Main Points

- Most studies were rated moderate risk of bias, while the majority of evidence was rated low strength or of insufficient strength to draw firm conclusions on comparative benefits and harms.
- ***Arthroplasty versus Anterior Cervical Discectomy and Fusion (ACDF):*** There was strong evidence that arthroplasty is associated with a substantially decreased likelihood of reoperation at 24 months with surgery at one level (SOE: High); there was low-strength evidence that 2-level arthroplasty versus ACDF is associated with decreased likelihood of reoperation at 24 months (SOE: Low), with similar results at longer followup times. There was moderate-strength evidence of no differences between arthroplasty and ACDF in improved postoperative pain or function with 1- level surgery (SOE: Moderate), whereas evidence was less strong with 2-level disease (SOE: Low) across various measures and timepoints.
- ***Anterior versus Posterior approach:*** There was low-strength evidence of similar reoperation rates in patients with radiculopathy and single-level disease (SOE: Low), but the likelihood of experiencing any serious adverse event was higher with posterior approaches than with ACDF in patients with 3 or more level disease (SOE: Low). There was inadequate evidence to determine comparative benefits between anterior and posterior approaches for other outcomes due to lower quality studies with inconsistent findings (SOE: Insufficient); additional studies are needed.
- ***Standalone cage versus Plate and cage in ACDF:*** There was no difference in fusion rates between the two fusion techniques at 12 months, with similar fusion rates at 24 and 36 months (SOE: Moderate). There were also no differences between a standalone cage versus plate and cage in postoperative improvement in arm pain, function, quality of life, or adjacent-level ossification following ACDF (SOE: Low). Few reoperations were reported.
- ***Laminoplasty versus laminectomy and fusion.*** There were little differences between surgical techniques in postoperative neurologic function (SOE: Moderate) or general function (SOE: Low), but the risk of experiencing a complication was lower with laminoplasty (SOE: Low), with no difference in reoperation rates (SOE: Moderate).
- ***Other comparisons.*** Evidence for other comparisons was limited. No studies meeting inclusion criteria were available to guide management of CDD in asymptomatic patients with radiographic spinal cord compression or to guide management of pseudarthrosis after anterior cervical fusion. Additional, well-conducted, randomized trials are needed.

Background and Purpose

This systematic review identifies and synthesizes research on treatments for cervical degenerative disc disease (CDD) in patients with or without cervical radiculopathy or myelopathy. This topic was nominated by the Congress of Neurological Surgeons (CNS), which published prior guidelines on the management of cervical degenerative disease in 2009.[1-3] This review is intended to be broadly useful to clinicians, patients, and policy makers, and will also inform the development of updated guidelines from CNS or others. This review included

nonoperative management of cervical degenerative disease as compared with operative management, which was not part of the previous CNS guidelines. Additionally, there were several gaps in the evidence identified in the previous CNS guidelines[3] that we addressed with this systematic review (e.g., the development of kyphotic deformity after surgery and its association with health outcomes; the effects of patient age, duration of symptoms, and MRI T2 hyperintensity as prognostic indicators; and the identification of optimal treatment for soft lateral cervical disc displacement causing radiculopathy).

Methods

This review follows standard methods for systematic reviews[4] that are further described in the full protocol available on the Agency for Healthcare Research and Quality (AHRQ) website: https://effectivehealthcare.ahrq.gov/products/cervical-degenerative-disease/protocol. The protocol was registered with PROSPERO (CRD42023386838). Searches were conducted in Ovid MEDLINE®, CINAHL®, EMBASE, and Cochrane CENTRAL databases from 1980 to July 1, 2022, and were supplemented by manual review of reference lists and a Federal Register Notice.

Investigators developed pre-established eligibility criteria defined by populations, interventions, comparators, outcomes, and setting in accordance with established methods[4] and revised the criteria with input from a technical expert panel and federal partners. The population included adults (≥18 years old) managed for symptomatic cervical degenerative disease (e.g., pain, radiculopathy, myelopathy) for all key questions (KQ), and patients with asymptomatic CDD for KQ1. For this review, management was defined as cervical spine surgery, non-surgical treatments, intraoperative monitoring, and pre- and post-operative imaging. Methods are discussed in more detail in the full report.

Results

A total of 4,471 references from electronic database searches and reference lists were reviewed. After dual review of titles and abstracts, 1,491 papers were selected for full-text review, of which 1,360 articles were excluded. Across all KQs, 106 studies in 131 publications on the comparative effectiveness and harms of management for cervical degenerative disease were included; 56 (in 80 publications) were RCTs and 49 (in 50 publications) were observational studies, and 1 was a systematic review. The largest number of studies evaluated the effectiveness of cervical arthroplasty compared to anterior cervical discectomy and fusion in patients with cervical spondylotic radiculopathy or myelopathy at one or two levels (KQ8, k=33), followed by comparative effectiveness and harms of surgery based on interbody graft material or device type in patients undergoing anterior cervical discectomy and fusion (KQ9, k=20). There was no evidence for comparative effectiveness and harms of surgery compared to non-operative treatment or no treatment (KQ1) or posterior approaches compared to revision anterior arthrodesis (KQ10). Main findings are summarized by Key Question in **Table 1**. Results are discussed in more detail in the full report.

Table 1. Summary of Findings: Cervical Degenerative Disease Treatment

Key Question	Comparison	Fusion *Effect Direction (SOE)*	Pain *Effect Direction (SOE)*	Function *Effect Direction (SOE)*	Quality of Life *Effect Direction (SOE)*	Adverse Events *Effect Direction (SOE)*
KQ 1. Radiographic and spinal cord compression and no myelopathy	Surgery vs. nonoperative treatment	No evidence	No evidence	No evidence	No evidence	No evidence
KQ 2. Radiographic spinal cord compression and mild to severe myelopathy	Surgery vs. nonoperative treatment	No evidence	No evidence	Insufficient evidence	No evidence	Insufficient evidence
KQ 3. In cervical degenerative disease	Surgery vs. nonoperative treatment	No evidence	Insufficient evidence	Insufficient evidence	No evidence	No evidence
KQ 4. In cervical degenerative disease	ACDF vs. ACDF + collar	Insufficient evidence	Insufficient evidence	Insufficient evidence	No evidence	No evidence
	ACDF vs. ACDF + electromagnetic stimulation (EMS)	Improved fusion rates favors EMS (SOE: Low)	Insufficient evidence	Insufficient evidence	No evidence	No evidence
	Laminoplasty vs. Laminoplasty + collar	Not applicable	No important difference (SOE: Low)	No important difference (SOE: Low)	No evidence	No evidence
	Laminoplasty vs. laminoplasty + exercise	Not applicable	Insufficient evidence	No evidence	No evidence	No evidence
KQ 5. In cervical radiculopathy	Anterior vs. posterior surgery	Insufficient evidence	<u>Neck and Arm pain</u>: No important difference (SOE: Low)	Insufficient evidence	Insufficient evidence	<u>Reoperation:</u> No important difference (SOE: Low)
KQ 6. In cervical degenerative disease with ≥3 level disease	Anterior vs. posterior surgery	Insufficient evidence	<u>Neck pain:</u> No important difference (SOE: Low) <u>Arm pain:</u> Insufficient evidence	No important difference (SOE: Low)	Insufficient evidence	<u>Mortality, severe dysphagia:</u> No important difference (SOE: Low) Reoperation (SOE: Insufficient) <u>Serious AE:</u> Moderate to Large favors anterior (SOE: Low)

Key Question	Comparison	Fusion *Effect Direction (SOE)*	Pain *Effect Direction (SOE)*	Function *Effect Direction (SOE)*	Quality of Life *Effect Direction (SOE)*	Adverse Events *Effect Direction (SOE)*
KQ7. In cervical myelopathy	Laminectomy vs. Laminoplasty and fusion	No evidence	Insufficient evidence	No important difference (SOE: Moderate)	No evidence	Reoperation: No important difference (SOE: Moderate) Adverse events: Moderate to Large favors laminoplasty (SOE: Low)
KQ8. In cervical radiculopathy and/or myelopathy	Arthroplasty vs. ACDF	Not applicable	No important difference (SOE: Moderate)	No important difference (SOE: Moderate)	No evidence	Reoperation: High favors arthroplasty (1-level SOE: High) (2-level SOE: Low) Serious AE: Small favors arthroplasty (SOE: Low) Neurological events: No important difference (1-level SOE: Low) (2-level SOE: Insufficient)
KQ9. In ACDF	Standalone cage vs. plate and cage	No important difference (SOE: Moderate)	Neck pain: No important difference (SOE: Low) Arm pain: Insufficient evidence	No important difference (SOE: Low)	No important difference (SOE: Low)	Adjacent level ossification: No important difference (SOE: Low)
	Titanium/titanium-coated vs. PEEK cage	Small favoring PEEK (SOE: Low)	Insufficient evidence	Small favoring PEEK (SOE: Low)	No evidence	Insufficient evidence
	Autograft vs. allograft vs. other osteogenic materials	Insufficient Evidence	Insufficient evidence	Insufficient evidence	Insufficient evidence	Adverse events: Large favors nonuse of BMP-2 (SOE: Low)
KQ 10. In pseudarthrosis after prior anterior fusion surgery	Posterior approach vs. revision anterior arthrodesis	No evidence	No evidence	No evidence	No evidence	No evidence

Key Question	Comparison	Fusion *Effect Direction (SOE)*	Pain *Effect Direction (SOE)*	Function *Effect Direction (SOE)*	Quality of Life *Effect Direction (SOE)*	Adverse Events *Effect Direction (SOE)*
KQ 11. In cervical myelopathy, prognostic utility of MRI for neurologic recovery	T2-weighted increased signal intensity and intensity ratio, sharp signal intensity	No evidence	No evidence	No evidence	No evidence	Neurologic recovery: favors no signal, less sharp signal, increased signal intensity ratio (SOE: Low)
KQ 12. Imaging to detect pseudarthrosis	Dynamic radiographs (asymptomatic population)	Predicts pseudarthrosis (SOE: Low)	Not applicable	Not applicable	Not applicable	No applicable
	Dynamic radiographs (symptomatic population)	Predicts pseudarthrosis (SOE: Low)	Not applicable	Not applicable	Not applicable	Not applicable
KQ 13. In cervical myelopathy	IONM vs. no IONM in ACDF	No evidence	No evidence	No evidence	No evidence	Neurologic complications: No important difference (SOE: Low)

Effect Direction: none, slight/small, moderate, or large effect/improvement
Strength of Evidence: low, moderate, high
none = no effect/no statistically significant effect

Strengths and Limitations

Most of the limitations of the evidence base are related to the rigor with which the studies were conducted, completeness of reporting key outcomes, and lack of comparative evidence. Limitations of these studies generally led to determination of insufficient evidence for many outcomes. Confounding by indication, lack of adequate control for confounding on important prognostic factors, as well as failure to adequately account for selection of patients and loss to follow-up were common methodologic concerns in NRSIs. Limitations of the review methods include limiting the evidence to English-language publications.

This review appears to provide the most comprehensive synthesis of evidence related to the comparative effectiveness of surgical treatment of CDD and identifies important gaps in the comparative evidence for many of them. Important strengths of this review include the use of a "best evidence" approach, where we focused our efforts on studies with least risk of bias, particularly randomized trials when available, and supplemented with nonrandomized studies that adjusted for potential prognostic variables where appropriate. Another strength is our focus on outcomes of primary importance to patients including pain, function, and quality of life, as improved patient outcomes may lead to higher quality patient care, as well as patient satisfaction with care. Additionally, interpretation of clinically important differences in mean change for continuous variables is challenging. Another strength of our review is our categorization of the magnitude of effects for function and pain outcomes using the system described in our previous reviews to facilitate interpretation of results across trials and interventions by providing a level of consistency and objective benchmarks for comparison. We also added two contextual questions (on the natural history of untreated spinal cord compression and on the prevalence of CDD with spinal cord compression in asymptomatic patients) to inform this review.

Future Research Needs

Research is needed to address gaps and deficiencies of existing studies. Additional research is needed to evaluate management of cervical degenerative disease that has not been addressed by existing studies, including well-designed prospective comparative NRSIs with protocols using methods for patient selection and treatment allocation that mitigate possible selection bias and imbalances in prognostic factors and that follow protocols established *a priori* for comparable treatment, evaluation, and measurement of groups.

Future trials should evaluate specific (or device-specific) harms and adverse events; ideally such studies would be powered to detect rare events. The large sample sizes available for administrative data may facilitate evaluation of rare outcomes and may demonstrate statistical significance when results may be of unclear clinical importance; many trials in this review had small sample sizes that precluded such analyses. In addition, trials should report the proportion of patients who experience a clinically important improvement in pain or function. To assess differential impact of patient characteristics and other factors, future trials should be adequately powered and include patients with diverse backgrounds, such as those who are disadvantaged due to socioeconomic factors, rural location, or geographic isolation; and from other underserved groups at risk for health disparities based on race, ethnicity, and disabilities. Studies should also estimate the minimally important between-group differences for included outcomes to facilitate interpretation of study findings.

Conclusions

There were few differences in benefits between surgical approaches and techniques compared in included studies for the treatment of cervical degenerative disease. However, there were some differences in the frequency of adverse events for some comparisons. There was substantial evidence that the risk of reoperation is much lower for artificial disc replacement than ACDF. Limited evidence also suggests a lower likelihood of experiencing any serious adverse event with ACDF than posterior cervical disc fusion, and a lower risk for any complication with laminoplasty compared with laminectomy and fusion. There was limited evidence on the role of nonoperative management instead of surgery or in addition to surgery to treat CDD, and no evidence to determine benefits and harms of a revision anterior arthrodesis or posterior approach in patients with pseudarthrosis after prior anterior cervical fusion.

References

1. Matz PG, Anderson PA, Kaiser MG, et al. Introduction and methodology: guidelines for the surgical management of cervical degenerative disease. J Neurosurg Spine. 2009 Aug;11(2):101-3. doi: /10.3171/2009.1.SPINE08712. PMID: 19769488.
2. Fehlings MG, Arvin B. Surgical management of cervical degenerative disease: the evidence related to indications, impact, and outcome. J Neurosurg Spine. 2009 Aug;11(2):97-100. doi: 10.3171/2009.5.SPINE09210. PMID: 19769487.
3. Congress of Neurological Surgeons. Guideline for the Surgical Management of Cervical Degenerative Disease. Congress of Neurological Surgeons; 2009. https://www.cns.org/guidelines/browse-guidelines-detail/surgical-management-of-cervical-degenerative-disea2022. Accessed January 27, 2023
4. Agency for Healthcare Research and Quality. Methods guide for effectiveness and comparative effectiveness reviews Agency for Healthcare Research and Quality. Rockville, MD: 2020. https://effectivehealthcare.ahrq.gov/products/collections/cer-methods-guide. Accessed January 27, 2023

1. Introduction

1.1 Background

The cervical spine is comprised of seven vertebrae with discs between the vertebrae that are comprised mostly of water. Cervical degenerative disease refers to a cascade of events that leads to changes of the vertebral discs resulting in disc desiccation and height loss. These changes may cause uncovertebral and facet joint hypertrophy (enlargement of vertebral joints) leading to vertebral foraminal narrowing (stenosis), which may cause radiculopathy (pain, numbness, and tingling radiating down the arm) as the exiting nerve roots are pinched, or more central stenosis with compression of the spinal cord and associated myelopathy (pain, numbness, and tingling due to spinal cord compression). Both cervical radiculopathy and cervical spondylotic myelopathy can sensory deficit and motor deficit, as well as pain. While both conditions can affect the neck and upper extremities, cervical spondylotic myelopathy can also cause poor proprioception and spasticity of the lower extremities resulting in gait disturbances, as well as disturbances in bladder function. Cervical radiculopathy and cervical spondylotic myelopathy may also exist simultaneously.

Although the etiology of cervical degenerative disease is not fully understood, it is a common condition that becomes more prevalent with age. The estimated prevalence of any spinal degenerative disease from 2005 to 2017, in people 65 and older, based on Medicare data of approximately 1.7 million individuals, is 27.3%, with the highest prevalence for degenerative disc disease (12.2%).[5] In a separate Medicare database study, 3,156,215 individuals were identified with degenerative cervical disease (incidence 18.9% for females, 13.1% for males between 2006 and 2012).[6] However, the presence of cervical degenerative disease may not correlate well with symptoms.[7] For example, one systematic review[8] found the prevalence of multilevel degenerative disc pathology to be 64.5% in asymptomatic subjects (compared with 89.7% in a symptomatic population).

1.1.1 Management of Cervical Degenerative Disease

Of the over 3 million individuals with cervical degenerative disease in the Medicare study mentioned above, 32% were treated nonoperatively and 7% were treated with spinal fusion (permanently joining two or more vertebrae) within a year of diagnosis.[6] Surgical treatment for cervical radiculopathy varies and includes both anterior and posterior based procedures. When approached anteriorly, intervertebral spacers and additional plating may be used, the vertebrae may or may not be fused, and the cervical disc(s) may or may not be replaced.[9] In addition to anterior cervical discectomy with fusion, cervical disc replacement and anterior cervical corpectomy (removal of the vertebral body) with fusion, surgical treatment for cervical spondylotic myelopathy also includes posterior based procedures; laminoplasty (surgery to enlarge spinal canal by cutting the bony roof [lamina] and allowing it to open like a door), laminectomy (surgery that enlarges spinal canal by removing a portion of the lamina), and laminectomy with fusion.[10] Nonoperative treatment of cervical degenerative disease includes analgesics, corticosteroids, neck immobilization, traction of the cervical spine, interventional approaches (e.g., radiofrequency ablation [a procedure that destroys nerve tissue that sends pain signals to the brain using radio waves), physical therapy, exercises, thermal therapy, and

avoidance of provocative activities.[11,12] The goals of both surgical and nonoperative treatments are to alleviate pain, improve neurologic function, and prevent progression or recurrence.

While cervical myelopathy and radiculopathy are clinical diagnoses, Magnetic Resonance Imaging (MRI) is used to confirm levels where compression of the spinal cord or nerve roots is evident. Various degenerative features can be seen on cervical MRI such as decreased vertebral height, disc height loss, osteophyte formation, disc bulging and location, hypertrophy and ossification of the posterior longitudinal ligament, spinal cord compression and flattening, and tethering (attachment) of the spinal cord to the spinal canal.[13] MRI findings can then help guide treatment. It is important to note that the presence of degenerative findings on MRI does not equate to symptomatic consequence. One study found that 28 percent of asymptomatic volunteers over the age of 40 years (N=23, levels=97) demonstrated cervical degenerative changes on MRI (versus 14 percent in those less than 40 years of age).[14] Intraoperative neuromonitoring (e.g., somatosensory, motor evoked potential measurements, spontaneous and triggered electromyography) is sometimes used during cervical spine surgery to provide intraoperative assessments of neural function and detect neurological injury during surgery to potentially mitigate or prevent further injury.

1.2 Purpose and Scope of the Review

This systematic review identifies and synthesizes research on treatments for cervical degenerative disease (CDD) in patients with or without cervical radiculopathy or myelopathy. This topic was nominated by the Congress of Neurological Surgeons (CNS), which published prior guidelines on the management of cervical degenerative disease in 2009.[1-3] This review is intended to be broadly useful to clinicians, patients, and policy makers, and will also inform the development of updated guidelines from CNS or others. This review also includes nonoperative management of cervical degenerative disease as compared with operative management, which was not part of the previous CNS guidelines.

2. Methods

2.1 Systematic Review Design Process

This Comparative Effectiveness Review (CER) follows methods of the Agency for Healthcare Research and Quality (AHRQ) Methods Guide for Effectiveness and Comparative Effectiveness Reviews (hereafter the "AHRQ Methods Guide").[4] All methods were determined a priori and a protocol was developed through a process that included collaboration with a technical expert panel (TEP), federal partners, and public input on key questions and study eligibility criteria. The protocol was registered on the PROSPERO systematic reviews registry (CRD42023386838) and published on the AHRQ web site: https://effectivehealthcare.ahrq.gov/products/cervical-degenerative-disease/protocol.

2.1.1 Key Questions

The review is defined by thirteen key questions that address the effectiveness and harms of treatments for CDD, as well as how effectiveness and harms may differ by patient and disease characteristics (e.g., age, gender, severity of disease, vertebral level(s) of involvement). Two contextual questions were also included to help inform the report. Contextual questions are not reviewed using systematic review methodology. The key questions, contextual question, and analytic framework (Figure 1) are below.

Key Question 1: In patients with radiographic spinal cord compression and no cervical spondylotic myelopathy, what are the comparative effectiveness and harms of surgery compared to non-operative treatment or no treatment?

Key Question 2: In patients with radiographic spinal cord compression and mild to severe myelopathy, what is the effectiveness and harms of surgery versus non-operative treatment or no treatment? How do the effectiveness and harms vary by level of severity of myelopathy at the time of surgery?

Key Question 3: In patients with cervical degenerative disease, what are the comparative effectiveness and harms of surgical compared to non-operative treatment?

Key Question 4: In patients with cervical degenerative disease, what are the comparative effectiveness and harms of therapies added on to surgery (pre- or post-operative) compared with the same surgery alone?

Key Question 5: In patients with cervical radiculopathy due to cervical degenerative disease, what are the comparative effectiveness and harms of posterior versus anterior surgery?

Key Question 6: In patients with cervical degenerative disease, what are the comparative effectiveness and harms of posterior versus anterior surgery in patients with greater than or equal to three level disease?

Key Question 7: In patients with cervical spondylotic myelopathy due to cervical degenerative disease, what are the comparative effectiveness and harms of cervical laminectomy and fusion compared to cervical laminoplasty in patients?

Key Question 8: In patients with cervical spondylotic radiculopathy or myelopathy at one or two levels, what are the comparative effectiveness and harms of cervical arthroplasty compared to anterior cervical discectomy and fusion?

Key Question 9: In patients undergoing anterior cervical discectomy and fusion, what are the comparative effectiveness and harms of surgery based on interbody graft material or device type?

Key Question 10: In patients with pseudarthrosis after prior anterior cervical fusion surgery, what are the comparative effectiveness and harms of posterior approaches compared to revision anterior arthrodesis?

Key Question 11: In patients with cervical spondylotic myelopathy, what is the prognostic utility of preoperative magnetic resonance imaging (MRI) findings for neurologic recovery after surgery?

Key Question 12: What is the sensitivity and specificity of imaging assessment for identifying symptomatic pseudarthrosis after prior cervical fusion surgery?

Key Question 13: In patients with cervical spondylotic myelopathy, what are the comparative effectiveness and harms of intraoperative neuromonitoring (e.g., with somatosensory or motor evoked potential measurements) versus no neuromonitoring on clinical outcomes in patients undergoing surgery?

For purposes of these key questions, we focused on symptomatic cervical degenerative disease; with the exception of Key Question 1, evaluation and management of asymptomatic disease is beyond the scope of this review.

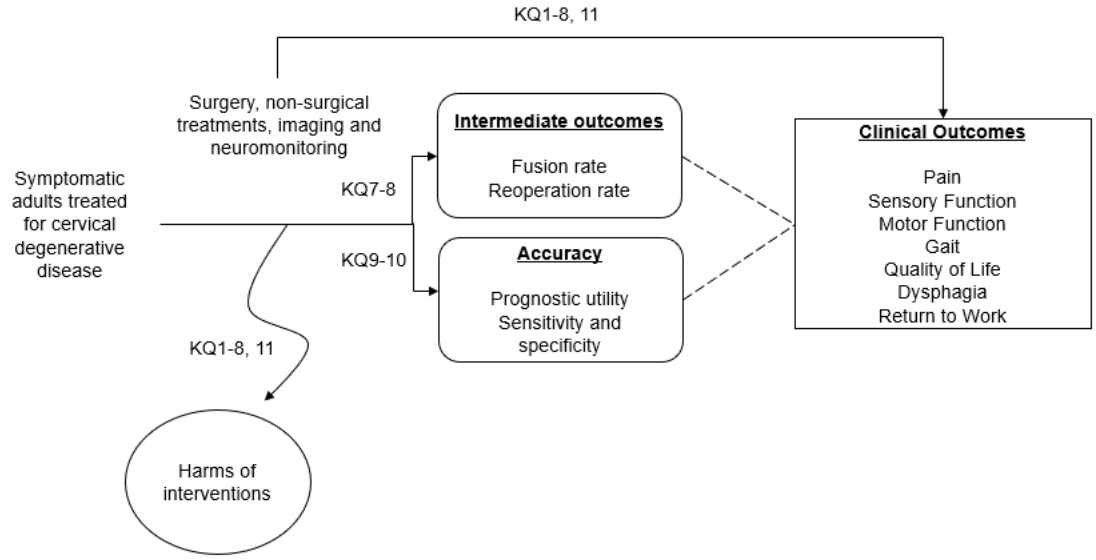
2.1.2 Contextual Questions

Contextual Question 1: What is the prevalence of cervical degenerative disease with spinal cord compression in asymptomatic patients?

Contextual Question 2: What is the natural history of untreated spinal cord compression in patients with cervical degenerative disease?

2.1.3 Analytic Framework

Figure 1. Analytic Framework



Abbreviations: KQ=key questions

The analytic framework illustrates how the populations, interventions, and outcomes relate to the KQ in the review.

2.2 Study Selection

2.2.1 Literature Search Strategy

We conducted electronic searches in Ovid MEDLINE®, EMBASE, and Cochrane CENTRAL from 1980 to July 1, 2022 (see **Appendix A1.1** for full strategies). For key questions that compare operative approaches, we searched databases for studies published after 2006 (studies published in 2007 or earlier were included in the 2009 guidelines).[3] Additionally, we reviewed all studies included in the 2009 guidelines for inclusion in this review.[3] For key questions not covered by the 2009 guidelines (e.g., operative versus nonoperative studies, neuromonitoring studies) we searched the databases from 1980 to the present in order to identify relevant, earlier studies based on when technologies such as neuromonitoring and advanced imaging were first used in research trials. Reference lists of included systematic reviews were screened for additional studies and relevant references were carried forward. A Federal Register notification for a Supplemental Evidence and Data for Systematic review (SEADS) portal was posted from August 12th to September 12th, 2022, for submission of unpublished studies. Searches will be updated while the draft report is posted for public comment.

2.2.2 Inclusion and Exclusion

Criteria were established *a priori* to determine eligibility for inclusion and exclusion of abstracts in accordance with the AHRQ Methods Guide.[4] The criteria for inclusion and exclusion of studies for this systematic review are based on the Key Questions and are described in **Table 1** (see **Appendix A** for complete details, and **Appendix B** for all included studies). More

information on data management methods can be found in **Appendix A2.1.** For studies meeting inclusion criteria, evidence tables were constructed, with results relevant to each KQ abstracted in **Appendix C**.

Table 1. PICOTS – Inclusion and Exclusion Criteria

	Include	Exclude
Population	• Age 18 and above with symptomatic cervical degenerative disease (e.g., pain, radiculopathy, myelopathy) for all KQs except for KQ1, which includes asymptomatic patients • Effectiveness and harms of surgery based on patient characteristics, disease characteristics and radiographic characteristics (e.g., age, gender, comorbidities [e.g., comorbid lumbar disease, autoimmune disease, neurological disease, mental illness, Down's syndrome], severity of cervical degenerative disease, Frailty Index, sagittal vertical aspect, degree of kyphosis, prior treatment [e.g., bracing, traction, medications, massage, acupuncture, injections, chiropractic care, spinal manipulation], duration of pain, skill of surgeon)	• Younger than 18 years • Patients without cervical degenerative disease • Nonhumans
Interventions	• Cervical spine surgery (e.g., discectomy, disc replacement, fusion up to T2, arthroplasty, laminectomy, laminoplasty, corpectomy, cervical hybrid surgery, foraminotomy, ACDF cage vs. ACDF cage + plate) • Non-surgical treatments (e.g., heat, exercise, acupuncture, drugs, radiofrequency ablation, steroid injections, Botox® for neck pain, psychological strategies [e.g., cognitive behavioral therapy], occupational therapy, multidisciplinary rehabilitation) • Intraoperative neuromonitoring • Imaging to identify symptomatic pseudarthrosis after cervical fusion surgery • Preoperative MRI to predict neurologic recovery in myelopathy	• Preoperative imaging using CT or plain films • KQ4: intraoperative therapy • KQ7: laminectomy without fusion
Comparators	• Any included intervention • Placebo, waitlist, active control • No comparator (KQs 11 and 12)	• Nonoperative intervention versus nonoperative intervention without surgical comparator
Outcomes	• Pain, sensory function, motor function, gait, quality of life (e.g., VAS, NRS, NDI, SF-36, SF-12, EQ-5Dm, mJOA score, Nurick score, MDI, PROMIS-29), dysphagia scales, return to work • Fusion rate, reoperation rate • Harms (e.g., withdrawals due to adverse events, serious adverse events, new symptomatic adjacent segment disease, postoperative infection, device failure, ossification of the posterior ligament, development of kyphotic deformity) • Sensitivity and specificity of imaging after cervical fusion surgery	• Nonvalidated instruments
Timing	• All time periods	
Setting	• Inpatient, outpatient, ambulatory surgical centers	

	Include	Exclude
Study types and designs	• RCTs, prospective trials and retrospective observational studies with a control group (study N≥50), current systematic reviews • KQs 11-12 and studies focused on harms as a primary outcome: large intervention series (N≥50; can be single arm, but everyone received the same intervention)	• KQ1-10: pre-post single-arm studies, case series (everyone selected based on outcome), case reports, systematic reviews published prior to 2007 • KQ11-12: pre-post non-intervention studies, case series, case reports, systematic reviews published prior to 2007
Language	• English language	• Non-English

Abbreviations: ACDF = anterior cervical discectomy and fusion; CT = computed tomography; EQ-5D = EuroQol-5 dimension instrument; KQ = key question; MDI = myelopathy disability index; MRI = magnetic resonance imaging; mJOA = modified Japanese orthopedic association scale; NDI = neck disability index; NRS = numerical pain rating scale; PROMIS-29 = patient reported outcome measurement information system; RCT = randomized controlled trial; QOL = quality of life; SF = short form health survey (12 or 36 items); VAS = visual analogue scale for pain

2.3 Risk of Bias Assessment of Individual Studies

Predefined criteria were used to assess the risk of bias (also referred to as quality or internal validity) for each individual included study, using criteria appropriate for the study design based on the AHRQ-EPC *Methods Guide for Effectiveness and Comparative Effectiveness Reviews,*[4] the Cochrane Back and Neck Group,[15] and the U.S. Preventive Services Task Force[16] (**Appendix D1.1**). Each study was independently reviewed for risk of bias by two team members. Any disagreements were resolved through consensus. Based on the risk of bias assessment, included studies were rated as having "low," "moderate," or "high" risk of bias. Studies rated high risk of bias were not excluded a priori, but were considered to be less reliable than low or moderate risk of bias studies when synthesizing the evidence. See **Appendix D1.1** for additional details.

Most studies were rated moderate risk of bias; we specifically call out in the text studies rated high risk of bias as extra caution should be exercised when drawing conclusions from such studies.

2.4 Data Analysis and Synthesis

Evidence tables identify study characteristics, results of interest, and risk of bias ratings for all included studies and summary tables highlight the main findings. Studies were reviewed and highlighted using a hierarchy-of-evidence approach, where the best evidence is the focus of the synthesis for each key question. Since the key questions varied in nature and scope, the approach to synthesis also varied. We analyzed the evidence according to KQ, using both qualitative (narrative) and where possible quantitative (meta-analysis) methods. Randomized controlled trials (RCTs) were prioritized and studies with lower risk of bias ratings were given more weight in our synthesis for each clinical indication and outcome.

Meta-analyses were conducted to obtain more precise effect estimates for comparative effectiveness of various interventions for cervical spine; analyses of randomized and nonrandomized evidence were conducted separately. A random effects model based on the profile likelihood method[17] was used to obtain pooled RR and MD. Statistical heterogeneity

among the studies was assessed using Cochran's χ^2 test and the I^2 statistic.[18] For analyses with at least 10 trials, we constructed funnel plots and performed the Egger test to detect small sample effects (a marker for potential publication bias).[19] All meta-analyses were conducted using Stata/SE 16.1 (StataCorp, College Station, TX). See **Appendix A2.1** for additional details on data synthesis and analysis.

To help determine the degree of effect, we examined the magnitude of relative risks and mean differences according to **Table 2**. There were instances where a statistically significant difference between treatments was of such a small magnitude as to not be clinically meaningful. Conversely, there were instances where a small, moderate, or large effect was found but was not statistically significant.

Table 2. Definition of effect sizes

Effect Size	Definition
Small effect	MD 0.5 to 1.0 points on a 0 to 10-point scale, 5 to 10 points on a 0 to 100-point scale SMD 0.2 to 0.5 RR/OR 1.2 to 1.4
Moderate effect	MD >1 to 2 points on a 0 to10-point scale, >10 to 20 points on a 0 to 100-point scale SMD >0.5 to 0.8 RR/OR 1.5 to 1.9
Large effect	MD >2 points on a 0 to10-point scale, >20 points on a 0 to 100-point scale SMD >0.8 RR/OR ≥2.0

Table 2 taken from the Cervical Degenerative Disease Protocol, published online at https://effectivehealthcare.ahrq.gov/sites/default/files/product/pdf/cervical-degenerative-protocol.pdf
MD = mean difference; OR = odds ratio; RR = relative risk; SMD = standardized mean difference

2.5 Grading the Strength of the Body of Evidence

The strength of evidence (SOE) for each body of evidence was assessed as high, moderate, low, or insufficient, using the approach described in the AHRQ Methods Guide,[4] based on study limitations, consistency, directness, precision, and reporting bias. These criteria were applied regardless of whether evidence was synthesized quantitatively or qualitatively. Strength of evidence ratings reflected our confidence or certainty in the findings. Strength of evidence was considered insufficient when evidence was lacking, sparse, or too conflicting such that we were unable to draw conclusions. SOE was initially assessed by one researcher and confirmed by a second. SOE was not conducted for composite outcomes. Descriptions of criteria and overall grades are described in full in **Appendix A** and **G**.

2.7 Peer Review and Public Commentary

An associate editor from a different EPC reviewed the draft report. Experts will be invited to provide external peer review of this systematic review; AHRQ will also provide comments. In addition, the draft report will be posted on the AHRQ website for 4 weeks for public comment. Comments will be reviewed and used to inform revisions to the draft report.

3. Results

3.1 Description of Included Studies

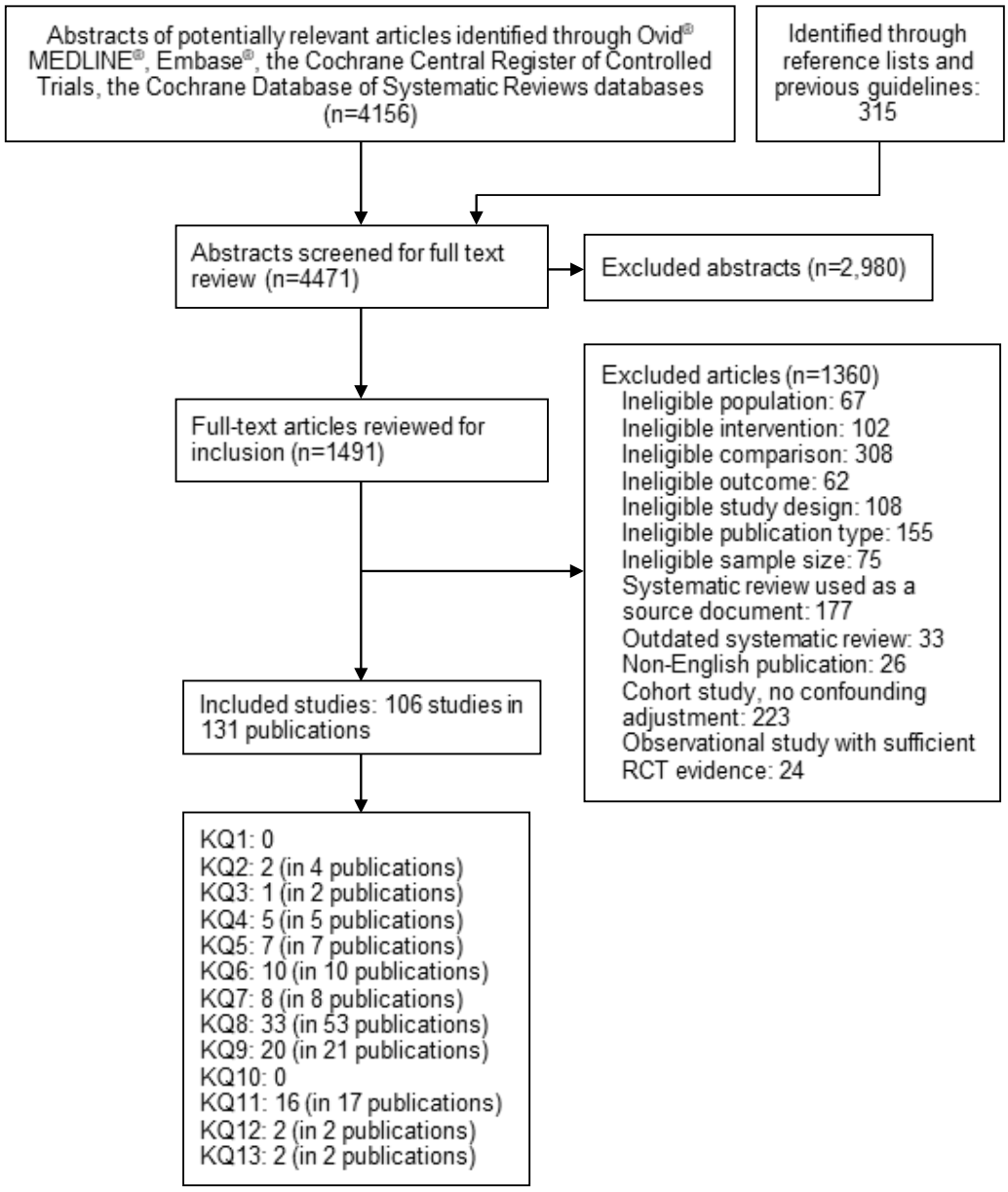
A total of 4,471 references from electronic database searches and reference lists were reviewed. After dual review of titles and abstracts, 1,491 papers were selected for full-text review, of which 1,360 articles were excluded. Of the 106 studies in 131 publications included across all key questions, 56 (in 80 publications) were RCTs and 49 (in 50 publications) were observational studies and one was a systematic review (**Figure 2**). Results are arranged by key question, then by outcome, and are summarized below, followed by tables in the accompanying text.

A list of excluded studies with reason for exclusion are in **Appendix E.** Data abstraction of study characteristics and results, quality assessment for all included studies, and details for grading SOE are available in **Appendices C, D, and G**, respectively.

Most studies were rated moderate risk of bias. For these studies we do not call out their risk of bias in the text. Instead we call out studies that were rated high risk of bias as additional caution should be exercised when interpreting study results.

Figure 2. Literature Flow Diagram

3. Results



Abbreviations: KQ = Key Question, RCT = randomized controlled trial

3.2 Key Question 1. In patients with radiographic spinal cord compression and no cervical spondylotic myelopathy, what are the comparative effectiveness and harms of surgery compared to non-operative treatment or no treatment?

No studies met eligibility criteria for Key Question 1.

3.3 Key Question 2. In patients with radiographic spinal cord compression and mild to severe myelopathy, what is the effectiveness and harms of surgery versus non-operative treatment or no treatment? How do the effectiveness and harms vary by level of severity of myelopathy at the time of surgery?

3.3.1 Key Findings

- Evidence from one small RCT and one small NRSI was inadequate to determine the benefits and harms of surgery versus conservative treatment for cervical myelopathy (SOE: Insufficient).

3.3.2 Description of Included Studies

One RCT (N=68) described in three publications[20-22] and one NRSI (N=80)[23] compared surgery versus conservative treatment for cervical myelopathy (**Appendix C**). The duration of followup in the randomized trial was 3 years[20,21] and 10 years.[22] The duration of followup in the NRSI was 3 years.[23] In the NRSI patients were stratified by degree of myelopathy (mild and moderate versus severe) in both the surgery and conservative treatment groups. In the RCT, all patients had slowly or non-progressing mild to moderate myelopathy. The RCT was conducted in the Czech Republic and received government funding; the NRSI was conducted in Italy and did not report funding.

The mean age of participants was 53 years with 29% females in the RCT and 66 years with 44% female in the NRSI. The duration of disease was 2 years (range 0.3 to 12 years) in the RCT and the mean duration of symptoms was 25 months (range 3 to 57 months) in the NRSI.

Surgery consisted of anterior decompression (N=22) with bone graft (N=20), corpectomy (N=6), and laminoplasty (N=5) in the RCT. An anterior approach was used in 1- or 2-level cord compression and a posterior approach was used in multilevel spinal stenosis. Surgery consisted of microsurgical anterior corpectomy, discectomy, use of titanium mesh and anterior plating in the NRSI. For 3- or multi-level corpectomy, posterior stabilization was also performed. Surgical patients wore a cervical collar for 4 weeks postoperatively. In the randomized trial, conservative treatment consisted of cervical collar, anti-inflammatory medication, and bed rest. However, surgical patients also received these treatments. Conservative treatment in the NRSI was similar to treatments in the RCT, but also included physiotherapy.

The RCT was rated moderate risk of bias due to lack of blinding and unclear randomization methods (**Appendix D**). The NRSI was rated high risk of bias due to unclear differences in patient baseline characteristics across groups and potential selection bias in treatment given (**Appendix D**). The strength of evidence for neurologic and general function was rated insufficient due to conflicting evidence from two small studies (**Appendix G**).

3.3.3 Detailed Analysis

3.3.3.1 Fusion

No studies reported fusion outcomes.

3.3.3.2 Pain

No studies reported pain outcomes.

3.3.3.3 Neurologic Function

Evidence from one small RCT and one small NRSI was inadequate to determine the benefits and harms of surgery versus conservative treatment on neurologic function in patients with cervical myelopathy (SOE: Insufficient).

In the randomized trial, patients were considered to be responders if mJOA scores (maximum 18 points) were improved or unchanged following treatment.[21] The likelihood of mJOA response was slightly less with surgery compared with conservative therapy at 6 months (N=66, 61% vs. 73%, RR 0.83, 95% CI 0.59 to 1.18) and at 36 months (N=59, 59% vs. 73%, RR 0.80, 95% CI 0.55 to 1.16), although differences were not statistically significant. However, mean mJOA scores were not different between surgery and conservative treatment at 6, 12, 24, and 36 months after controlling for baseline values. Ten-year followup of the randomized trial (N=47) also found no differences between treatment groups on the mJOA (14 vs. 15, p=0.114).[22]

In the NRSI, patients were divided into four groups (N=20 patients per group) and followed for 3 years: patients with mild to moderate myelopathy treated with surgery; patients with mild to moderate myelopathy treated conservatively; patients with severe myelopathy treated with surgery; patients with severe myelopathy treated conservatively.[23] Mild to moderate myelopathy was defined as a mJOA score of 12 and above, severe myelopathy as a score below 12. Patients with severe myelopathy experienced a longer duration of symptoms (40 months) than patients with mild to moderate disease (10 months) and were more likely to receive multilevel surgery than surgical patients with mild to moderate disease. Mean mJOA scores improved over time for both surgery and conservative treatment but favored surgery at 12 and 36 months in patients with mild to moderate myelopathy (12 months mJOA: 15.4 vs. 14.2, p=0.03; 36 months: 16.1 vs. 15.2, p=0.013). In patients with severe myelopathy improvement in mJOA scores was greater with surgery compared with conservative treatment beginning at 6 months (6 months mJOA: 9.5 vs. 7.9, p=0.045; 12 months: 11.5 vs. 8.6, p=0.001; 36 months: 12.45 vs. 8.65, p<0.001).

3.3.3.4 General Function

Evidence from one small RCT and one small NRSI was inadequate to determine the benefits and harms of surgery versus conservative treatment on general function in patients with cervical myelopathy (SOE: Insufficient).

The time required to complete the 10-meter Walk Test in the randomized trial (N=66) increased over time through 24 months in patients treated with surgery (baseline: 7.9 seconds; 6 months: 8.7 sec; 12 months: 9.9 sec; 24 months: 11.7 sec; 36 months: 9.4 sec), whereas there was little change in time needed to complete the 10-meter walk throughout the followup period with conservative treatment (baseline: 7.4 sec; 6 months: 7.2 sec; 12 months: 7.4 sec; 24 and 36 months: 7.5 sec).[20] These differences in walk time between treatments were statistically significant (p-value range 0.034 to 0.003), although the differences between groups is not likely clinically meaningful. Ten-year followup of the RCT (N=47) found no differences on the 10-meter Walk Test (7.3 seconds vs. 7.1 seconds, p=0.207).[22] There was no difference, however, in the NRSI, between treatment with surgery versus conservative therapy on the 10-meter Walk Test in patients with mild to moderate myelopathy, whereas there was greater improvement on the 10-Meter Walk Test with surgery in patients with severe myelopathy at 12 and 36 months (12

months: 11.4 seconds vs. 14.4 seconds, p=0.005; 36 months: 10.30 seconds vs. 14.10 seconds, p=0.002).[23]

In the RCT, patients were videoed performing activities of daily living (ADL) such as buttoning a shirt, brushing teeth and hair, walking, going up and down stairs, and running and were evaluated by blinded observers on a 7-point improvement scale that ranged from 3 (excellent) to -3 (poor); 0 represented no change in ability.[20] Patients treated with surgery showed a greater likelihood of improvement in ADLs compared with conservative treatment at 6 months (20% vs. 5.9%) but there was also a greater likelihood of worsening in ADLs with surgery (20% vs. 8.8%) at 6 months. There were no differences between treatments in changes in ADL abilities at 12, 24, or 36 months. Video evaluation of decreased ability to perform ADLs was also not different between treatment groups at 10 years (mean of two evaluators: 56.8% vs. 50%, p>0.05).[22] However, with the limited sample size available, this 10-year followup was likely underpowered to demonstrate a difference between surgery and conservative treatment.

Although more patients in the RCT reported that their disease course had improved after surgery compared with conservative therapy at 6 months posttreatment (61% vs. 20%, p=0.001), self-perception of improved diseased course deteriorated over time in the surgery group (p=0.019 for negative trend) and was 20% at 36 months compared with a relatively stable course with conservative treatment.[20] Ten-year followup of the randomized trial (N=47) found no difference between treatment groups on a subjective evaluation of worsened status (45.5% vs. 56%, p=0.47).[22]

The physical component summary and the mental component summary score on the SF-12 were not different posttreatment (unclear posttreatment time) in patients with mild to moderate myelopathy who received surgery compared with patients who received conservative therapy (PCS: 37.4 vs. 37.95, p=0.75; MCS: 47.5 vs. 46.7, p=0.78).[23] However, improvement in scores was greater with surgery versus conservative treatment in patients with severe myelopathy (PCS: 53.3 vs. 26.85, p<0.001; MCS: 61.2 vs. 31.4, p<0.001).

3.3.3.5 Quality of Life

No studies reported quality of life outcomes.

3.3.3.6 Harms

The NRSI reported that two patients with severe myelopathy who received conservative treatment demonstrated progressive neurological worsening (defined as a worsening of 1 point on the mJOA).[23] Surgical complications in this study included 5/40 patients (1.25%) who experienced airway obstruction, graft displacement, and/or wound hematoma. There were no deaths.

The findings of the NRSI, particularly the findings in patients with severe myelopathy, should be interpreted with caution as the individuals in the severe myelopathy group who received conservative treatment consisted of those who refused surgery against medical advice, which may have introduced selection bias.

3.4 Key Question 3. In patients with cervical degenerative disease, what are the comparative effectiveness and harms of surgical compared to non-operative treatment?

3.4.1 Key Findings

- There was inadequate evidence from one small RCT on the comparative effectiveness of ACDF, physiotherapy, and treatment with a cervical collar on pain and function in patients with cervico-brachial pain without spinal cord compression (SOE: Insufficient).

3.4.2 Description of Included Studies

One RCT (N=81) described in two publications[24,25] compared treatment for cervico-brachial pain with cervical decompression and fusion, physiotherapy, or neck collar (**Appendix C**). All patients had nerve root compression on MRI without spinal cord compression, a history of pain for 3 or more months and were followed for 16 months. The study was conducted in Sweden.

The mean age of participants was 47 years and 46% were female; race or ethnicity were not reported. The worst affected level was C5-C6 (49%) followed by C6-C7 (37%). Prior treatments included physiotherapy (85%; physiotherapy uses a hands-on approach to healing, e.g., massage, fascial releases, whereas physical therapy uses hands-on methods but also incorporates physical exercises and use of a cervical collar (42%). Mean duration of pain was 34 months (range 5 to 120 months).

Surgery consisted of ACDF using the Cloward technique and fusion achieved with purified cow bone graft; one patient received a posterior laminectomy. Surgical patients sometimes wore a collar for 1 to 2 days postoperatively. Physiotherapy included traction (70%), strengthening exercises (56%), stretching exercises (56%), massage (33%), heat (33%), and transcutaneous electrical stimulation (22%), among other modalities. Patients treated with cervical collars used a rigid collar during the day and an optional soft collar at night for 3 months.

The trial was rated moderate risk of bias due to lack of blinding and overlap in treatments after 16 weeks (**Appendix D**). The strength of evidence for pain, neurologic function and general function was rated insufficient due to limited evidence from one small trial (**Appendix G**).

3.4.3 Detailed Analysis

3.4.3.1 Fusion

No studies reported fusion outcomes.

3.4.3.2 Pain

There was inadequate evidence from one small RCT on the comparative effectiveness of ADCF, physiotherapy and treatment with a cervical collar on pain in patients with cervico-brachila pain without spinal cord compression (SOE: Insufficient).

There were no differences between treatments in current pain or worst pain using the VAS (0-100) at baseline.[24] At 14 to 16 weeks followup patients treated with surgery experienced less "current" pain that patients treated with a collar (N=54, 0-100 VAS: 27 vs. 48, $p<0.01$), but there was no difference between surgery, physiotherapy, and use of a collar in "current" pain at 16 months (N=81, VAS: 30 vs. 39 vs. 35, $p>0.05$). Results were similar regarding "worst" pain with

surgical patients experiencing less "worst" pain than collar patients at 14-16 weeks (N=54, VAS: 43 vs. 64, p<0.001) but no differences in "worst" pain between treatments at 16 months (N=81, VAS: 42 vs. 53 vs. 52, p>0.05, respectively).

3.4.3.3 Function

3.4.3.3.1 Neurological Function

There was inadequate evidence from one small RCT on the comparative effectiveness of ADCF, physiotherapy and treatment with a cervical collar on neurologic function in patients with cervico-brachila pain without spinal cord compression (SOE: Insufficient).

Specific muscle strength before and after treatment was also assessed.[25] Patients in the surgery group experienced greater improvements in muscle strength (strength expressed as the ratio of the affected to the unaffected side) at 14-16 weeks in pinch grip, elbow extension and shoulder internal rotation compared with patients treated with physiotherapy and greater improvements in wrist flexion and elbow flexion compared to those treated with a cervical collar (data not provided). At 16 months, patients treated with surgery experienced greater improvements in wrist extension, elbow extension, shoulder abduction, and shoulder internal rotation compared with patients treated with physiotherapy. There were no differences in strength improvement between surgery and collar treatment or between physiotherapy and collar treatment at 16 months (data not provided).

At 14 to 16 weeks posttreatment, there was no difference in the likelihood of improvement in paresthesias with surgery compared with physiotherapy or collar treatment (N=81, 52% vs. 45% vs. 37%, p>0.05) but a large increase in the likelihood of improvement in sensory loss with surgery compared with either treatment (41% vs. 15%, RR 2.75, 95% CI 1.0 to 7.5, both comparisons with surgery).[25] At 16 months, there remained no difference between treatment in the likelihood of improvement in paresthesias between surgery, physiotherapy, and treatment with a collar (N=81, 58% vs. 67% vs. 66%, p>0.05). There was also no difference between treatments in the likelihood of improvement in sensory loss at 16 months (N=81, 27% vs. 14% vs. 15%, p>0.05).

3.4.3.3.2 General Function

There was inadequate evidence from one small RCT on the comparative effectiveness of ADCF, physiotherapy and treatment with a cervical collar on general function in patients with cervico-brachila pain without spinal cord compression (SOE: Insufficient).

The ability to complete basic activities of daily life (e.g., dressing, prolonged sitting) to more rigorous physical activity (e.g., running, heavy work) was assessed using the disability rating index (DRI).[24] Overall mean score on the DRI ranges from 0 to 100, with ability on each of 12 activities rated using a 0-100 VAS scale indicating "without difficulty" to "not at all." There was no difference between treatment with surgery versus physiotherapy at 14-16 weeks on improvement in disability, however treatment with surgery resulted in improved dressing and heavy work compared with treatment with a collar, while treatment with physiotherapy was associated with greater ability to walk, sit for a long time, and complete heavy work compared with collar treatment (p<0.05, data not provided). At 16 months the ability to do heavy work was greater with surgery compared to the other treatments (p<0.05, data not provided). No other differences on the DRI were noted.

3. Results

Although findings from this small study tended to favor surgery, especially in the short term, these findings should be interpreted with caution due to patients receiving additional treatments beyond the randomized treatment and the heterogeneity of treatment (especially physiotherapy). After 16 weeks, 8/27 surgery patients (30%) underwent a second surgery. Additionally, one patient treated with physiotherapy (4%) and five treated with collar (19%) underwent surgery. Forty-one percent of surgery patients (11/27) received physiotherapy as did 44% (12/27) of patients treated with a collar. Additionally, the use of specific physiotherapy modalities (e.g., traction, exercises, cryotherapy) varied and was at the discretion of the local physiotherapist.

3.4.3.4 Quality of Life

This study did not report quality of life outcomes.

3.4.3.5 Harms

This study did not report harms or adverse events.

3.5 Key Question 4. In patients with cervical degenerative disease, what are the comparative effectiveness and harms of therapies added on to surgery (pre- or post-operative) compared with the same surgery alone?

3.5.1 Key Findings

- Laminoplasty
  - There was low strength evidence of no difference in pain and function between use of a post-operative collar plus laminoplasty versus laminoplasty alone (SOE: Low).
  - There was inadequate evidence to determine the effects on pain with laminoplasty plus exercise versus laminoplasty alone (SOE: Insufficient).
- ACDF
  - There was low-strength evidence that use of post-operative pulsed electro-magnetic field (PEMF) stimulation after ACDF was associated with increased fusion versus treatment with ACDF alone (SOE: Low); pain and function were similar with or without PEMF after ACDF (SOE: Low).
  - There was inadequate evidence to determine the effects on fusion, pain, and function of ACDF plus post-operative collar compared with ACDF alone (SOE: Insufficient).

3.5.2 Description of Included Studies

Five RCTs (N=546)[26-30] compared surgery plus post-operative therapy to surgery alone (**Appendix C**). The average mean followup duration was 12 months (range 1 week to 2 years). Two trials were conducted in Japan,[29,30] and 1 trial each in the U.S.,[28] Sweden,[26] and China.[27]

The average study mean age of participants was 59 years (range 47 to 73 years); the average proportion of females in studies was 38% (range 29% to 47%). Two trials reported race, one enrolling a majority of White participants (93%)[28] and the other enrolling Chinese participants.[27] Studies enrolled patients with clinical and/or radiological evidence of cervical myelopathy[27,29,30] or radiculopathy.[26,28] Patients had 1-2 level disease in 1 trial (N=33),[26] 1-4 levels (60% had 2 levels) in 1 trial (N=323),[28] and a mean of 4.5 levels in 1 trial (N=90).[29] Two trials did not report number of disease levels.[27,30]

One trial was rated low risk of bias,[27,28] and the remainder were rated moderate risk of bias (**Appendix D**). Methodological limitations included unclear blinding of providers or assessors and high loss to followup. Evidence for pain and function with laminoplasty plus exercise versus laminoplasty alone and evidence for fusion, pain and function for ACDF plus post-operative collar versus ACDF alone were rated insufficient due to limited evidence from one small trial each (**Appendix G**).

3.5.3 Detailed Analysis

3.5.3.1 Laminoplasty Plus Nonoperative Therapy Versus Laminoplasty

Three RCTs (N=190) assessed laminoplasty plus post-operative Philadelphia collars[27,29] or exercise therapy incorporating 3 months of daily strengthening and range of motion exercises.[30]

3.5.3.1.1 Fusion

No study reported fusion outcomes.

3.5.3.1.2 Pain

There was no difference in pain between the use of a post-operative collar plus laminoplasty versus laminoplasty alone. (SOE: Low). There was inadequate evidence to determine the effects on pain with laminoplasty plus exercise versus laminoplasty alone (SOE: Insufficient).

Single-door laminoplasty plus rigid Philadelphia collar worn for 3 weeks post-operatively was associated with less improvement in mean VAS scores (0-10 scale) than laminoplasty alone at weeks 1 (0.8 vs. 3.8, p=0.023) and 2 (-0.9 vs. 1.8, p=0.046) in one trial rated low risk of bias (N=35), with no difference at other timepoints (3 weeks: -1.2 vs. 1.1, p=0.148) or at other followup times (6 weeks and 3, 6, and 12).[27] One trial (N=90) compared modified double-door laminoplasty plus Philadelphia collar worn for 2 weeks post-operatively and found no differences in change in VAS (0-10 scale) at 12 months (-0.19 vs. -0.04, p>0.05) or throughout the study period (p=0.487).[29]

One RCT (N=65) found no difference in mean VAS scores (0-100 scale) for neck pain and stiffness at 2 weeks and 3 months postoperative between muscle-preserving laminoplasty with exercises versus laminoplasty alone (3 months: -1.8 vs. -2.5, p=0.623).[30]

3.5.3.1.3 Function

3.5.3.1.3.1 Neurologic Function

There was no difference in neurologic function between the use of a post-operative collar plus laminoplasty versus laminoplasty alone. (SOE: Low).

One trial of open-door laminoplasty (N=35) found no difference on mJOA scores between 3 weeks of post-operative collar versus no collar at 6 weeks (mJOA: 13.8 vs. 13.3, p=0.613)[27] or longer followup. This was consistent with 12-month results from the second collar trial (N=90) which reported no difference in end-of-study mJOA scores between 2 weeks of post-operative collar use and no collar (11.1 vs. 11.8, p=0.22).[29]

3.5.3.1.3.2 General Function

There was no difference in general function between the use of a post-operative collar plus laminoplasty versus laminoplasty alone. (SOE: Low). Two trials (N=125) of laminoplasty with or without the addition of a postoperative Philadelphia collar for 2 or 3 weeks were consistent in finding no difference in SF-36 PCS and MCS scores with collar use compared to no collar. One RCT (N=35) of single-door laminoplasty found no differences in SF-36 scores between the use of a post-operative collar for 3 weeks versus no collar at 6 weeks after surgery when controlling for baseline scores (PCS: 6.4 vs. 2.8; MCS: 4.1 vs. 0, p>0.05) or at longer followup times (3, 6, 12, 24 months).[27] One RCT (N=90) of double-door laminoplasty plus 2 weeks of postoperative collar use versus no collar also found no difference at 12 months in change in SF-36 PCS or MCS scores (PCS: 1.5 vs. 1.4, p>0.05; MCS: 0.1 vs. 0.4, p>0.05).[29] The trial of open-door laminoplasty also found no difference on NDI between 3 weeks of post-operative collar and no collar at 6 weeks (NDI: 24.8 vs. 34.0, p=0.147) or at longer followup.[27]

3.5.3.1.4 Quality of Life

No study reported quality of life outcomes.

3.5.3.1.5 Harms

No study reported harms or adverse events.

3.5.3.2 ACDF Plus Nonoperative Therapy Versus ACDF

One trial (N=33) assessed ACDF versus ACDF plus rigid Philadelphia collar worn for 6 weeks postoperative[26] and one trial (N=323) compared ACDF with ACDF plus pulsed electromagnetic field stimulation (PEFM), delivered using a Cervical-Stim device for 4 hours daily from 1 week to 3 months postoperatively in a trial of active smokers (all patients wore a cervical collar for 1 week postoperatively).[28]

3.5.3.2.1 Fusion

There was inadequate evidence to determine the effects on fusion between ACDF with or without collar use (SOE: Insufficient). Use of post-operative PEMF stimulation after ACDF was associated with increased fusion versus treatment with ACDF alone (SOE: Low).

All ACDF patients in one 24-month trial (N=33) achieved radiographic fusion regardless of collar use (100% vs. 100%).[26] Surgical details were not provided.

PEFM was associated with improved fusion at 6 months in one trial (N=323) based on a per protocol analysis versus ACDF with no PEFM (N=240; 83.6% vs. 68.6%, p=0.0065); fusion rates were also improved in ITT analyses assuming missing patients fused (N=323; 85.9% vs. 76.3%, p=0.0269) or imputing patient status at last visit (N=281; 78.2% vs. 64.8%, p=0.0127), but not when assuming missing patients did not fuse (65.6% vs. 56.3%, p=0.0835).[28] However, there was no difference in fusion rates in the per protocol analysis at 12 months.[28] This study used a Smith-Robinson technique with allograft and cervical plate system.

3.5.3.2.2 Pain

The ACDF trial of PEFM versus no PEFM found similar VAS scores for shoulder/arm pain at rest or with activity at 6 and 12 months postoperative (date provided in graph form)[28] (SOE: Low).

3.5.3.2.3 Function

3.5.3.2.3.1 General Function

There was inadequate evidence to determine the effect on general function of ACDF plus post-operative collar compared with ACDF alone for all time points (SOE: Insufficient).

Collar use was associated with greater improvement in SF-36 PCS scores from baseline than ACDF without a collar at 6 weeks (MD 5.8; 95% CI 0.8 to 10.7), 3 months (MD 6.8; 95% CI 0.4 to 13.1), 6 months (MD 7.4; 95% CI 1.4 to 13.4), and 12 months (MD 7.5; 95% 0.3 to 14.6), but not at 24 months (MD 4.9; 95% CI -0.8 to 10.5; p=0.088).[26] In the same trial, there was no difference in mean change in SF-36 MCS scores at 6 weeks (MD -1.9; 95% CI -11.1 to 7.4) or at longer postoperative followup times.[26]

Six-weeks' collar use was associated with greater improvement in NDI scores from baseline than no collar at 6 weeks (MD -4.4; 95% CI -8.6 to -0.2), but not at 3 months (MD -2.1, 95% CI -8.0 to 3.8) or at other timepoints.[26] There was no difference in NDI scores between daily PEFM and no stimulation at 6 months (31.0 vs. 23.0, p>0.05) or 12 months' postoperative (25.6 vs. 22.8, p>0.05).[28]

3.5.3.2.4 Quality of Life

No study reported quality of life outcomes.

3.5.3.2.5 Harms

No study reported harms or adverse events.

3.6 Key Question 5. In patients with cervical radiculopathy due to cervical degenerative disease, what are the comparative effectiveness and harms of posterior versus anterior surgery?

3.6.1 Key Findings

- There was low-strength evidence of no differences in neck and arm pain between anterior versus posterior approaches short term (3, 6 months) and intermediate term (12, 24 months) (SOE: Low).
- There was inadequate evidence to determine benefits of anterior versus posterior approaches for neck pain (immediately postoperative), fusion, neurologic function, general function, or quality of life (SOE: Insufficient).
- There was low-strength evidence of no difference between approaches in the likelihood of reoperation (SOE: Low).
- Neurologic deficits were reported inconsistently and various measures were used across studies, however there was low-strength evidence of no differences between approaches were reported (SOE: Low).
- No serious adverse events with either approach were reported in RCTs; evidence on specific adverse events was limited (SOE: Insufficient).
- One NRSI (N=46,598) reported higher 30-day mortality with ACDF versus posterior cervical foraminotomy (PCF), but there were very few deaths, providing inadequate evidence of any difference between approaches on mortality (SOE: Insufficient).

3.6.2 Description of Included Studies

Three RCTs (N=277)[31-33] compared anterior versus posterior approaches (**Appendix C**). The average mean followup duration was 33 months (range 14 to 60 months). One trial was conducted in the U.S.,[33] one in Germany,[32] and one in Egypt.[31] All three trials were conducted at single sites. The average study mean age of participants for the trials was 44 years (range 43 to 44 years); the average proportion of females in trials was 56% (range 50% to 66%). No trials reported race. All three trials limited enrollment to patients with radiculopathy; two trials excluded patients with myelopathy,[31,33] and the other did not report myelopathy.[32] Patients in all three trials had single-level disease. One trial was rated moderate risk of bias[33] and two trials were rated high risk of bias (**Appendix D**). [31,32] One trial stated that no funding was received,[32] and two trials did not address funding.[31,33] Primary methodologic concerns were unclear randomization and treatment allocation concealment, dissimilarity between treatment groups at baseline and lack of assessor blinding.

Four retrospective NRSIs (N=47,684), including one database study, compared anterior versus posterior procedures (**Appendix C**).[34-37] Three NRSIs were conducted in the U.S.[34,35,37] and one in the U.K.[36] Three studies[34-36] drew patients from a single site and one[37] used an insurance administrative database (N=46,598). The average study mean age of participants was 50 years (range 48 to 53 years); the average proportion of females in studies was 44% (range 31% to 54%). One study reported race, enrolling a majority of White participants (88%).[35] All four NRSIs limited enrollment to patients with radiculopathy. Patients had single-level disease in three NRSIs.[34,36,37] A mean of 2.6 surgical levels was reported in one study.[35] Funding was not

reported in two NRSIs,[34,36], one was government funded[37]and one stated that no funding was received. [35] Three NRSIs were rated moderate risk of bias[35-37] and one was rated high risk of bias (**Appendix D**).[34]Common methodologic limitations were unclear loss to follow-up and lack of clarity regarding assessor blinding. Additionally, lack of clarity regarding patient enrollment and comparability of treatment groups at baseline combined with inadequate adjustment for confounding for prognostic variables were concerns resulting in the NRSI being rated high risk of bias.

For many outcomes, authors did not provide adequate data to calculate effect sizes and confidence intervals. Although NRSI may have adjusted for some outcomes, authors did not always provide adjusted estimates for our outcomes of interest.

Evidence was insufficient for fusion, neurologic function, general function, quality of life, mortality and serious adverse events, based on a combination of two or more of the following: high risk of bias, inconsistent findings, and lack of precision (**Appendix G**).

3.6.3 Detailed Analysis

3.6.3.1 Anterior versus Posterior

The anterior approach used was anterior cervical foraminotomy (ACF) in one RCT,[31], anterior cervical decompression without fusion (ACD) in one RCT[33] and anterior cervical decompression and fusion (ACDF) in two RCTs[32,33] and all four NRSIs.[34-37] All studies used posterior cervical foraminotomy as the comparator.

3.6.3.1.1 Fusion

There was inadequate evidence to determine benefits and harms of anterior versus posterior surgical approaches on cervical fusion (SOE: Insufficient).

One RCT (N= 30) rated high risk of bias reported that no participants in either the anterior cervical foraminotomy group or the posterior cervical foraminotomy group had radiologic evidence of instability on cervical x-rays at time of discharge or at a mean of 14 months.[31]Authors did not define stability or criteria for determining fusion.

3.6.3.1.2 Pain

There were no differences in neck and arm pain between anterior versus posterior approaches in the short (3, 6 months) and intermediate term (12, 24 months) (SOE: Low); there was inadequate evidence to determine the benefits and harms of anterior versus posterior approaches on neck pain immediately post-operative (SOE: Insufficient).

One small trial (N=30) rated high risk of bias reported that anterior cervical foraminotomy (ACF) was associated with lower neck pain VAS scores (0-10 scale) within a week of discharge ($p<0.001$), however the reported confidence interval for the difference between groups suggested no difference (MD -3.13, 95% CI -4.52 to 1.74) and may be a typographical error.[31] A larger RCT (N=175) also rated high risk of bias, compared ACDF versus PCF at 3, 6, 12, and 24 months for arm pain VAS (0-100 scale), neck pain VAS (0-100 scale) and NASS pain (0-6 scale).[32] The mean differences across measures did not change with time and there were no differences between ACDF and PCF in arm pain VAS (range from -1 to 1), neck pain VAS scores (range from 1 to 4) or NASS pain scores (range from -0.1 to 0.1) at any timepoint. Statistical tests were not reported and reported data were inadequate to calculate confidence intervals for effect sizes, but the authors noted that the clinical results were the same in both

groups. The third RCT (N=72) rated moderate risk of bias, reported similar rates of patient-reported complete or partial pain improvement (unvalidated measure) for anterior approaches (ACD and ACDF) versus PCF at day 1 postoperatively (100% vs. 100%, RR 1.00), at 2 months (98% vs. 100%, RR 0.98, 95% CI 0.94 to 1.02, p=0.32), and at approximately 60 months postoperatively (96.5% vs. 100%, RR 0.96, 95% CI 0.90 to 1.03, p=0.32).[33]

Findings for pain from two NRSIs were consistent with those of the RCTs. The larger study (N=688) found no difference in mean scores for VAS arm pain (0-10 scale) at 3 months (4.20 vs. 3.82, MD 0.38, p>0.05), 12 months (4.06 vs. 4.07, MD 0.01, p>0.05) or 24 months (3.85 vs. 4.48, MD -0.63, p>0.05).[36] In the smaller NRSI (N=70) rated high risk of bias, there were no differences between ACDF versus PCF in VAS score (0-10 scale, not specified for arm or neck pain, 2.6 vs. 3.0, MD -0.4, p =0.04) at 12 months.[34] Reported estimates appear to be unadjusted.

3.6.3.1.3 Function

3.6.3.1.3.1 Neurologic Function

There was inadequate evidence to determine benefits and harms of anterior versus posterior approaches on neurologic function for all time points (SOE: Insufficient).

One RCT (N=175) rated high risk of bias[32] reported similar mean NASS neurology scores (0-6 scale) for ACDF and PCF and that no patient had deterioration of symptoms. Means were consistent at 3, 6, 12, and 24 months (range MD -0.2 to 0.2). Statistical tests were not reported and data were inadequate to calculate confidence intervals, but the authors noted that the clinical results were the same in both groups.

3.6.3.1.3.2 General Function

There was inadequate evidence to determine benefits of anterior versus posterior approaches on general function for all timepoints and measures (SOE: Insufficient).

There was no difference in function between ACF and PCF at a mean of 14 months reported by one RCT (N=30)[31] rated high risk of bias based on Odom's criteria: Excellent (73% vs. 60%, RR 1.22, 95% CI 0.73 to 2.04, p=0.44) or Good (20% vs. 33%, RR 0.60, 95% CI 0.17 to 2.07, p=0.42) or in the proportion of patients with functional outcome rated as satisfactory (0% for both) or poor (6.7% for both). One NRSI (N=688), reported no difference between ACD and ACDF on the Core Outcome Measures Index-neck (COMI-neck, 0-10 scale), which has items for pain, function, symptom-specific well-being, quality of life and disability. [36] Mean changes in COMI-neck scores (0-10 scale) were similar at 3 months (-2.38 vs. -2.31, p=0.88) and 6 months (-2.94 vs. -2.67, p=0.55); at 24 months the mean COMI-neck scores were also similar (4.16 vs. 4.72, p>0.05; mean change not reported). The proportion of patients who achieved minimum clinically important difference on the COMI-neck score (decrease ≥2 points) was also similar at 3 months (50% vs. 56%, RR 0.89, 95% CI 0.65 to 1.24), 12 months (59% vs. 58%, RR 1.02, 95% CI 0.76 to 1.36), and 24 months (57% vs. 50%, RR 1.14, 95% CI 0.71 to 1.83). One NRSI (N=70) rated high risk of bias found no difference between ACDF versus PCF in Pain Disability Questionnaire (PDQ) functional status subscale scores (0 to 90 scale, (31.3 vs. 43.2, MD – -11.9, p=0.30) or PDQ total score (52.8 vs. 69.6, p=0.50)[34] One RCT (N=175) rated high risk of bias reported Hilibrand criteria ratings (Poor, Satisfactory, Good, Excellent, measure not validated) for ACDF versus PCF at 3, 6, 12 and 24 months.[32] Data were not available to calculate effect sizes, but the authors noted that the clinical results were the same in both groups at all timepoints: Excellent (84% vs. 83% at 3 months, and 76% vs. 79% at 24 months).

3.6.3.1.4 Quality of Life

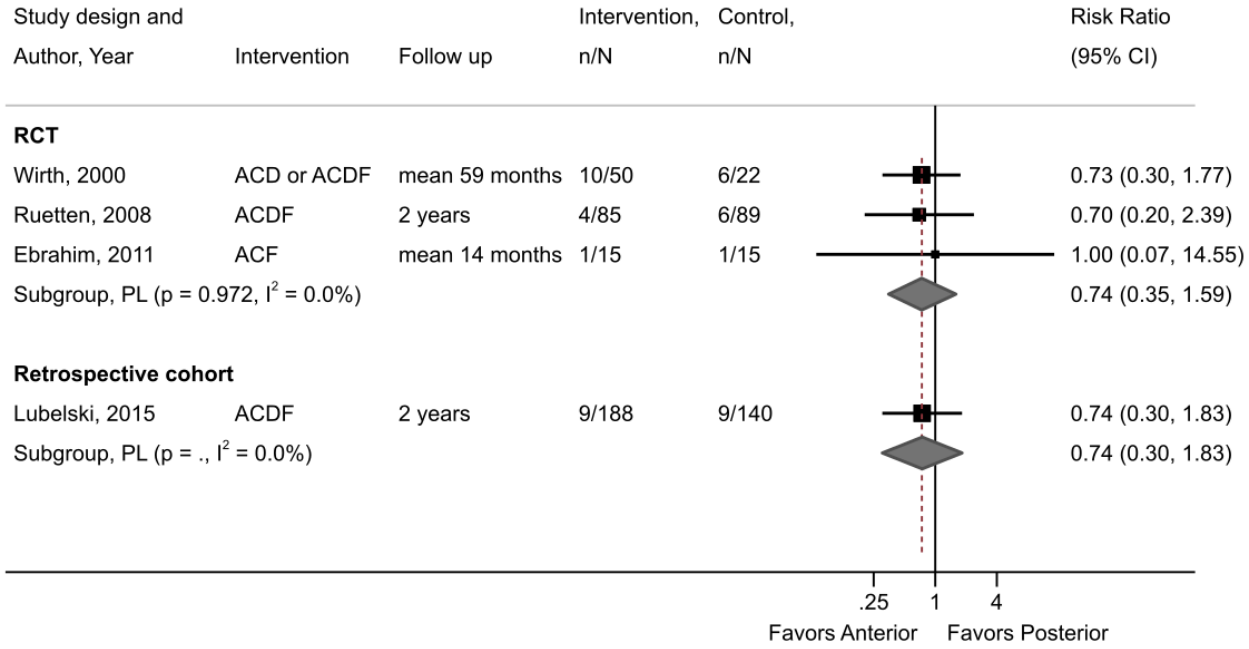
There was inadequate evidence to determine the benefits and harms of anterior versus posterior approaches at all time points (SOE: Insufficient).

One NRSI (N=70) rated high risk of bias found no difference in EuroQOL-5 Dimensions (EQ-5D, scale 0-1) at 12 months for ACDF (MD 0.69, 95% CI 0.61 to 0.77) versus PCF (MD 0.72, 95% CI 0.64 to 0.80, p=0.60).[34]

3.6.3.1.5 Reoperation

There were no differences in the likelihood of reoperation between anterior and posterior procedures across three RCTs[31-33] (2 of which were rated high risk of bias) or in one retrospective NRSI (N=328)[35] and effect estimates were similar (SOE: Low). (**Figure 3**)

Figure 3. Reoperation: Anterior versus posterior procedures



ACF = anterior cervical foraminotomy, ACD = anterior cervical decompression without fusion; ACDF = anterior cervical discectomy and fusion; CI = confidence interval; PL = profile likelihood.

3.6.3.1.6 Harms

There were no differences in neurologic deficits between anterior and posterior approaches, although results were reported inconsistently (SOE: Low); Reporting of other adverse events was limited (SOE: Insufficient).

Description of and reporting on serious adverse events was limited. All three RCTS (2 of which were rated high risk of bias) reported that no serious adverse events occurred for any patients.[31-33] One RCT (N=72) that compared ACD and ACDF to PCF reported zero deaths.[33] One propensity score matched NRSI (N=46,598) reported higher 30-day mortality with ACDF versus PCF (MD 1 event per 10,000 cases, 95% CI 0.0 to 2.0 per 10,000 cases, p=0.012).[37] Although the MD is significant, it is small, suggesting the possibility of 0 to 2 deaths with PCF. Given that administrative data are subject to misclassification and potential for inadequate adjustment for confounders, this finding should be interpreted cautiously.

3. Results

Neurologic deficits were reported inconsistently across studies. One RCT (N=72) found no difference in anterior versus posterior approaches for new weakness (8% vs. 14%, RR 0.59, 95% CI 0.14 to 2.40, p=0.46) or new numbness (6% vs. 9%, RR 0.66, 95% CI 0.12 to 3.68, p=0.63).[33] The other two RCTs reported specific neurologic deficits: in one small trial (N=30) no patients in either group developed Horner's syndrome;[31] the other trial (N=175) reported that no patients experienced damage to myelin resulting in paralysis of any degree.[32] One NRSI (N=70) reported that one patient who underwent PCF experienced C6 nerve injury, but did not provide data for patients who underwent ACDF.[34] Central nervous system complications at 30 days postoperatively was similar between anterior and posterior procedures in a large NRSI (N=46,598, MD 4 per 10,000, 95% CI -14 to 22 per 10,000).[37]

Dysphagia was reported inconsistently across studies. One RCT (N=175) reported transient difficulty swallowing for three patients who underwent ACDF and no patients who underwent PCF.[32] In a propensity score matched NRSI (N=46,598), ACDF was associated with higher rates of dysphagia/dysphonia at 30 days versus PCF (MD 14.5 per 1,000 cases, 95% CI 12.6 to 16.4 per 1000, p<0.001).[37] Neither study provided information on severity of dysphagia or need for intervention.

One large NRSI (N=46,598) reported that the following were rare but more common with ACDF versus PCF within 30 days after surgery: vascular injury (MD 2 per 10,000 cases, 95% CI 1 to 3 per 10,000 cases, p=0.001), CSF leak (MD 2 per 10,000 cases, 95% CI 1 to 3 per 10,000 patients, p= 0.002) and deep venous thrombus (9 per 10,000 cases, 95%CI 2 to 16 per 10,000 patients, p = 0.01). There were no differences between anterior and posterior approaches for pulmonary embolism (MD 2 per 10,000, 95% CI -9 to 12 per 10,000 cases).[37]

3.7 Key Question 6. In patients with cervical degenerative disease, what are the comparative effectiveness and harms of posterior versus anterior surgery in patients with greater than or equal to three level disease?

3.7.1 Key Findings

- There was low-strength evidence of no difference in neck pain, neurologic function and general function intermediate term (12 to 15 months) for ACDF versus posterior cervical decompression and fusion (PCDF) or laminoplasty for three or more levels (SOE: Low).
- The evidence for fusion, neck pain (short term), arm pain, neurologic function (short term) and quality of life was inadequate to draw conclusions (SOE: Insufficient).
- There was inadequate evidence to draw conclusions on reoperation rates between ACDF and posterior procedures (SOE: Insufficient).
- There was low-strength evidence that mortality and severe dysphagia did not differ between ACDF and laminoplasty or PCDF (SOE: Low).
- Rates of new neurologic complications and serious adverse events were inconsistently reported across studies and rare in general; there was low-strength evidence that posterior approaches were more commonly associated with a moderate to large increase in the odds of experiencing a neurologic adverse event and serious adverse event compared with ACDF (SOE: Low).

3.7.2 Description of Included Studies

One RCT[38] and nine NRSIs[39-47] compared anterior (i.e., ACDF) versus posterior surgery (i.e., laminoplasty, PCDF) at three or more levels for treatment of CDD **(Appendices C-D)**.

The RCT (N=34)[38] compared ACDF with posterior laminoplasty for participants with CSM (71%) or ossification of the posterior longitudinal ligament (OPLL) (29%) involving three (71%) or four (29%) levels. Fewer participants randomized to ACDF were diagnosed with OPLL (24% vs. 35%), had four-level disease (18% vs. 41%) or were smokers (12% vs. 41%). Mean participant age was 62 years and 26 percent were female.[38] Race/ethnicity was not reported. Average follow-up time was 41 months. This trial was conducted in China and was rated high risk of bias.

Across the nine NRSIs, one prospective[42] and eight retrospective,[39-41,43-47] sample sizes ranged from 245 to 13,884 (total N=41,982). The average study patient age was 61 years (range 54 to 63 years) and 43 percent were female (range 31% to 52%). Three studies reported race/ethnicity (White: range 65.5% to 82.3%; Black: 12.3% to 17.0%; Hispanic: 0.5%; Other: 17.7% to 19.1%).[39,45,46] The anterior approach was ACDF (with or without corpectomy) in all nine studies[39-47] and also included anterior cervical corpectomy and fusion (ACCF) in one study.[41] The posterior approach was posterior cervical discectomy and fusion (PCDF) in six studies,[39,40,42-44,46], laminectomy and fusion in two studies[41,45] and laminoplasty in two studies.[45,47] Two studies included three treatment groups; one with two anterior arms[41] and one with two posterior arms.[45] The number of involved levels varied across the studies but most included three to five levels; one study included only three levels[46] and another only four levels.[43] One NRSI was rated low risk of bias[44] and the remainder were rated moderate risk of bias.[39-43,45-47]

Evidence was insufficient for fusion, pain (short and long term), neurologic function (short term), quality of life, and reoperation based on a combination of two or more of the following: high risk of bias, inconsistent findings, and lack of precision (**Appendix G**).

3.7.3 Detailed Analysis

3.7.3.1 Fusion

There was inadequate evidence to determine the benefits and harms of anterior versus posterior surgical approaches on fusion in participants with three or more level disease (SOE: Insufficient).

One retrospective NRSI that used propensity score matching (N=12,248) found that PCDF was associated with substantially higher odds of pseudarthrosis at 12 months compared with ACDF (OR 2.43, 95% CI 1.96 to 3.01) at three levels.[46] The randomized trial did not report fusion.

3.7.3.2 Pain

There was low-strength evidence of no difference in neck pain in the intermediate term (SOE: Low); there was inadequate evidence for neck pain in the short term and arm pain in the intermediate term in participants with three or more level disease (SOE: Insufficient).

One RCT (N=32) rated high risk of bias reported no differences between 3- or 4-level ACDF and laminoplasty in neck pain scores (VAS, 0-10 scale) at 3 months (MD -0.10, 95% CI -0.46 to 0.26) and 6 months (MD 0, 95% CI -0.18 to 0.18) or at 12 months (MD 0.10, 95% CI -0.23 to 0.43) and 15 months (MD -0.10, 95% CI -0.44 to 0.24).[38] Similarly, there were no differences between ACDF (with and without corpectomy) and PCDF at three to five levels for NRS (0-10) neck pain scores (median 2 vs. 2, adjusted OR 0.67, 95% CI 0.37 to 1.21) or arm pain scores (median 1 vs. 0.5, adjusted OR 0.99, 95% CI 0.51 to 1.93) at 12 months in one retrospective NRSI (N=245).[39]

3.7.3.3 Function

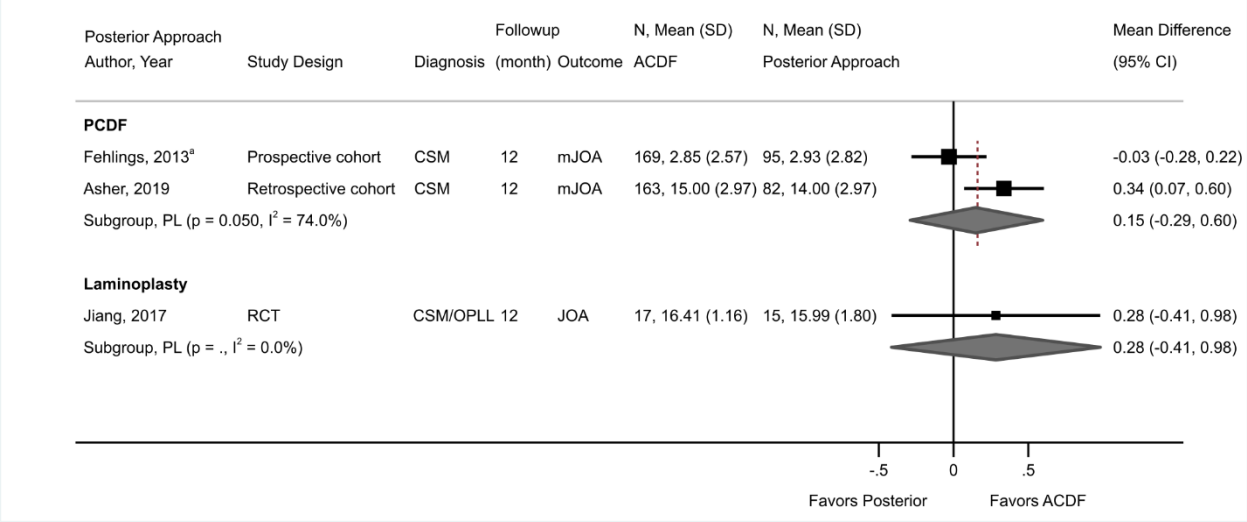
3.7.3.3.1 Neurologic Function

There was low-strength evidence of no difference in neurologic function between anterior and posterior approaches in participants with three or more level disease in the intermediate term (SOE: Low); there was inadequate evidence for determining the benefits and harms on neurologic function in the short term (SOE: Insufficient).

There was no difference in neurologic function at intermediate term (12 months) in one small RCT rated high risk of bias (N=32, MD 0.28, 95% CI -0.41 to 0.98, JOA scores, 0-18 scale)[38] and two NRSIs rated moderate risk of bias (N=506, MD 0.15, 95% CI -0.29 to 0.60, I^2=74.0%, mJOA scores, 0-18 scale)[38,39,42] that compared ACDF with posterior laminoplasty (RCT) or PCDF (NRSIs) for 3- to 5-level disease (**Figure 4**). (SOE: Low) There was also no difference between groups in JOA scores short term in the RCT (N=32): 3 months (MD -0.40, 95% CI -1.76 to 0.96) and 6 months (MD 0.20, 95% CI -1.14 to 1.54).[38] The pooled estimate across the two NRSIs had substantial heterogeneity (**Figure 4**), which may be due in part to different study designs, variables controlled for in multivariate analyses, and types of posterior procedures used. The prospective NRSI[42] showed no difference between groups and included patients who underwent laminoplasty (14%) (all others had PCDF); it was unclear which baseline confounders

were controlled for in this study. The retrospective NRSI[39] showed a large improvement with ACDF versus PCDF approaches; multivariate logistic regression models controlled for 19 different baseline variables.

Figure 4. Neurologic function (JOA or mJOA scores): Anterior versus posterior approaches for ≥3 levels



CI = confidence interval; CSM = cervical spondylotic myelopathy; JOA = Japanese Orthopaedic Association; mJOA = modified Japanese Orthopaedic Association; OPLL = ossification of the posterior longitudinal ligament, PCDF = posterior cervical decompression and fusion; PL = profile likelihood; RCT = randomized controlled trial; SD = standard deviation.

[a] Posterior approach included laminoplasty (14% of patients) or laminectomy and fusion (86% of patients)

One prospective NRSI (N=264) assessed neurologic function with the Nurick score (0-5 scale) and found no difference between 3- to 5-level ACDF and posterior approaches (laminectomy and fusion [86%] or laminoplasty [14%]) in mean change from baseline to 12 months after adjusting for baseline characteristics (MD in change scores 0.19, 95% CI -0.20 to 0.58[42]).

3.7.3.3.2 General Function

There were no differences between anterior and posterior surgery for 3- to 5-level disease at intermediate term (12 months) for any function measure reported across two NRSIs (N=509)[39,42] (SOE: Low). One prospective NRSI (N=264) compared ACDF with laminectomy and fusion (86%) or laminoplasty (14%) and reported the change in NDI scores compared with baseline (MD in change scores -0.97, 95% CI -7.15 to 5.21, scale unclear), SF-36 PCS scores (MD in change scores -1.90, 95% CI -5.30 to 1.50, 0-100 scale) and SF-36 MCS scores (MD in change scores 0.42, 95% CI -2.30 to 3.14, 0-100 scale).[42] One retrospective NRSI (N=245) compared ACDF (with and without corpectomy) with PCDF and reported median NDI scores (16 vs. 17, adjusted OR 0.76, 95% CI 0.42 to 1.37).[39]

3.7.3.4 Quality of Life

There was inadequate evidence to determine the benefits and harms of anterior versus posterior approaches on quality of life in participants with three or more level disease (SOE: Insufficient).

3. Results

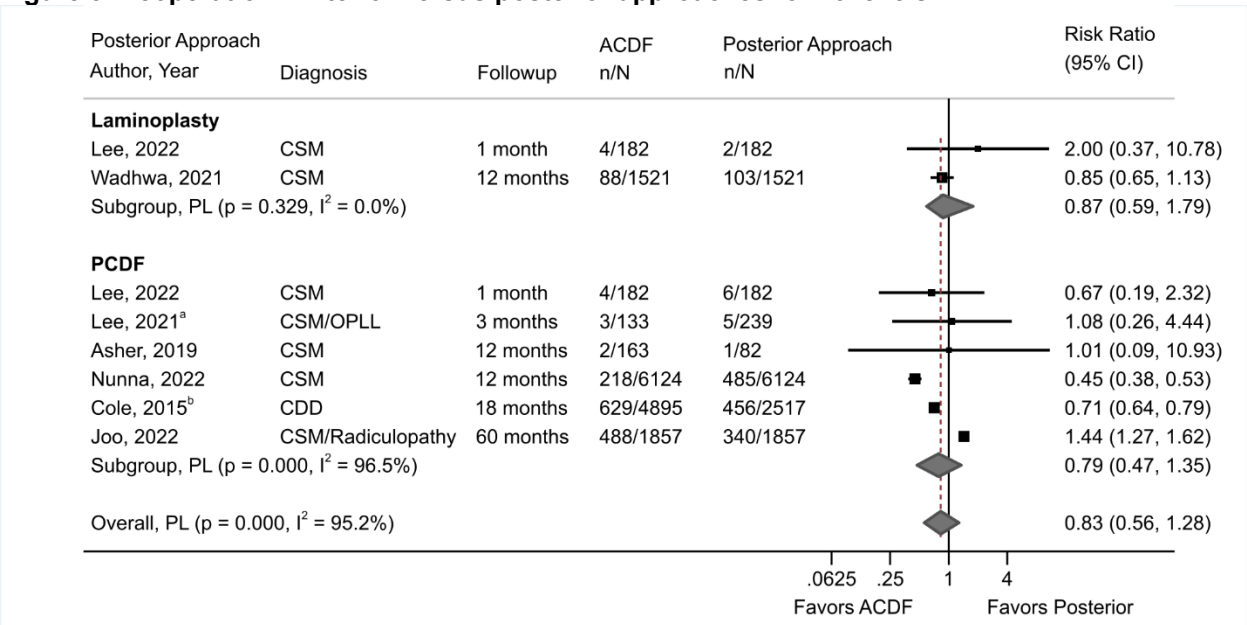
One retrospective cohort study (N=245) found no difference between 3- to 5-level ACDF (with and without corpectomy) and PCDF in EQ-5D scores intermediate term at 12 months (aOR 1.36, 95% CI 0.76 to 2.44, referent = ACDF) after adjusting for a number of baseline variables.[39]

3.7.3.5 Reoperation

There was inadequate evidence to draw conclusion on reoperation rates between ACDF and posterior procedures (SOE: Insufficient).

Seven NRSIs (N=27,579) that compared ACDF with posterior procedures at three or more levels reported reoperation/revision rates.[39,41,43-47] In pooled analysis at any timepoint based on longest follow-up (range 1 to 60 months), there were no differences between ACDF versus laminoplasty (2 NRSIs, N=3,406, 5.4% vs. 6.2%, RR 0.87, 95% CI 0.59 to 1.79, I^2=0%)[45,47] or versus PCDF (6 NRSIs, N=24,355, 10.1% vs. 11.8%, RR 0.79, 95% CI 0.47 to 1.35, I^2=96.5%);[39,41,43-46] however, heterogeneity was substantial for the latter comparison (**Figure 5**). Exclusion of one outlier study[43] at 60 months that included patients with both myelopathy and radiculopathy reduced heterogeneity slightly and resulted in a moderate reduction in the likelihood of reoperation for ACDF compared with PCDF at any timepoint (1-18 months, 5 NRSIs, N=20,641, 7.4% vs. 10.4%, RR 0.59, 95% CI 0.42 to 0.95, I^2=82.4%)[39,41,44-46] These results were driven by two large administrative data studies.[41,46] There was no difference between ACDF and PCDF at 1 to 3 months (2 NRSIs, N=736, RR 0.82, 95% CI 0.32 to 2.08, I^2=0%).[44,45] ACDF was associated with a higher risk of reoperation compared with PCDF (N=3,714, RR 1.44, 95% CI 1.27 to 1.62) in one study at 60 months.[43] It is challenging to draw firm conclusions from this data as definitions of reoperation and revision varied or were not specified across the studies, there were differences in posterior approach used, and the pooled estimates were mainly driven by two large administrative data studies.

Figure 5. Reoperation: Anterior versus posterior approaches for ≥3 levels



CI = confidence interval; CSM = cervical spondylotic myelopathy; OPLL = ossification of the posterior longitudinal ligament, PCDF = posterior cervical decompression and fusion; PL = profile likelihood.

[a] Study included patients with myelopathy and OPLL

[b] Anterior approach included ACDF or ACCF

One large NRSI (N=12,248) that used administrative data and propensity score matching reported reoperation outcomes that could not be included in the meta-analysis.[46] PCDF was associated with substantially higher odds of wound-specific revision surgery at 1 month (1.2% vs. 0.4%, OR 3.02, 95% CI 2.56 to 3.49) and moderately lower odds of additional anterior or posterior fusion at 12 months (4.3% vs. 7.0%, OR 0.60, 95% CI 0.44 to 0.76) compared with ACDF at three levels.

3.7.3.6 Harms

3.7.3.6.1 Neurologic Deficits

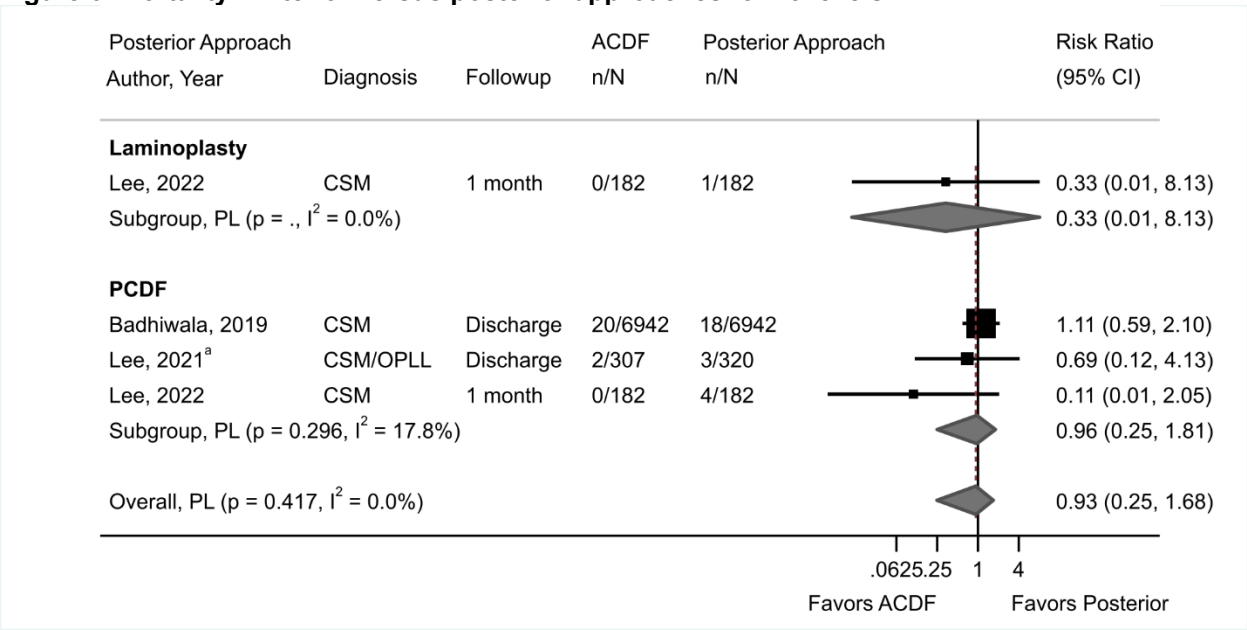
There was low-strength evidence that posterior approaches were more likely associated with a moderate to large increase in the odds of experiencing a neurologic adverse event compared with ACDF (SOE: Low). Reporting of neurological events varied across one RCT (N=32)[38] and six NRSIs (total N=37,095, range 245 to 13,884).[39-42,46,47] The RCT reported no cases of postoperative worsening of myelopathy or C5 root palsy with either 3- or 4-level ACDF versus posterior laminoplasty.[38] Central nervous system complications (not further defined) were rare through 90 days after ACDF (<0.7%) and posterior laminoplasty (0.9%) at three or more levels in one NRSI (N=3,042).[47] Two NRSIs reported that PCDF was associated with moderately higher odds of "neurological complications" compared with ACDF at three or more levels but did not provide further details: 0.59% vs. 0.35% (adjusted OR 1.7, 95% CI 1.0 to 2.8) immediately postoperative in one study (N=13,884)[40] and 1.8% vs. 1.1% (OR 1.6, 95% CI 1.08 to 2.38) at 1 month in another (N=7,412).[41] Two other NRSIs reported no difference between ACDF and PCDF at three to five levels in new neurological deficits (N=264, 4.1% vs. 3.2%, RR 1.31, 95% CI 0.35 to 4.95)[42] or new motor deficits (N=245, 2% vs. 0%)[39] at 12 months. One large NRSI (N=12,248) reported no difference between PCDF and ACDF in the incidence of postoperative coma (0.4% vs. 0.6%, OR 1.26, 95% CI 0.75 to 1.77).[46]

3.7.3.6.2 Mortality

There was low-strength evidence that mortality did not differ between ACDF and laminoplasty or PCDF (SOE: Low).

Three NRSIs (total N=15,057, range 546 to 13,884) that compared anterior with posterior approaches at three or more levels found no difference in short-term mortality after ACDF versus posterior laminoplasty at 1 month (1 NRSI, N=364, 0% vs. 0.05%, RR 0.33, 95% CI 0.01 to 8.13)[45] and ACDF versus PCDF at hospital discharge to 1 month (3 NRSIs, N=14,875, 0.3% vs. 0.3%, RR 0.96, 95% CI 0.25 to 1.81, I^2=17.8%)[40,44,45] (**Figure 6**). One NRSI (N=12,248) reported no deaths in either arm (ACDF vs. PCDF) and was unable to be included in the pooled analysis.[46]

Figure 6. Mortality: Anterior versus posterior approaches for ≥3 levels



CI = confidence interval; CSM = cervical spondylotic myelopathy; OPLL = ossification of the posterior longitudinal ligament, PCDF = posterior cervical decompression and fusion; PL = profile likelihood.

[a] Study included patients with myelopathy and OPLL

3.7.3.6.3 Dysphagia

There was low-strength evidence that the likelihood of experiencing severe dysphagia did not differ between ACDF and laminoplasty or PCDF (SOE: Low).

Severe dysphagia was rare across two NRSIs that compared ACDF with PCDF or posterior laminoplasty. There were two cases (1%) requiring a nasogastric tube in one study (N=245)[39] and one case (0.5%) requiring an unplanned readmission 11 days postsurgery in the other (N=364);[45] all three cases occurred in the ACDF arms. (SOE: Low)

One RCT (N=32)[38] and seven NRSIs (total N=41,172, range 245 to 13,884)[39-41,43,44,46,47] also reported dysphagia but did not report the severity; frequencies ranged from 2.7% to 14.0% after ACDF and from 0% to 3.6% after PCDF across six NRSIs (N=38,130)[39-41,43,44,46] most of which reported a substantial to moderate decrease in the odds/risk of dysphagia with PCDF (OR range 0.20 to 0.61), and from <0.7% to 5.9% versus 0% to <0.7% in the ACDF versus laminoplasty arms, respectively, across one small RCT (N=32)[38] and one large NRSI (N=3,042) with no differences between treatments.[47]

3.7.3.6.4 Serious Adverse Events

There was low-strength evidence that posterior approaches were more likely associated with a moderate to large increase in the odds of experiencing a serious adverse event compared with ACDF (SOE: Low).

One RCT (N=32) reported that intraoperative dural tear occurred in 5.9% of ACDF versus 11.8% of PCDF patients (RR 0.50, 95% CI 0.05 to 5.01) and that there were no cases of instrumentation failure or malposition, infection or hematoma.[38]

Across the NRSIs, reporting of serious adverse events varied and were generally rare and generally occurred more often with posterior approaches versus ACDF.

3. Results

Thrombolic events were rare across eight NRSIs (total N=41,718, range 245 to 13,884) with followup immediately postoperative to 12 months.[39-41,43-47] The frequency of DVT or PE ranged from 0% to 2.3% (ACDF) versus 0% to 4.3% (PCDF or posterior laminoplasty). Four of the studies (N=37,258) reported that posterior approaches were associated with moderate to large increases in the odds of experiencing a thrombolic event compared with ACDF (range of ORs 1.75 to 3.7).[40,41,43,46]

Stroke/cerebrovascular events occurred variably across three NRSIs with short term followup (1 to 3 months); one study (N=546) reported no events in either arm (ACDF vs. PCDF or posterior laminoplasty),[45] one study (N=627) reported more events after ACDF (1.8% vs. 0% PCDF, p=0.016)[44] while the third found that PCDF was associated with a large increase in the odds of stroke compared with ACDF (N=12,248, 4.2% vs. 2.5%, OR 1.68, 95% CI 1.48 to 1.89).[46]

Sepsis was rare across three NRSIs (total N=7,302, range 546 to 3,714).[43,45,47] One study reported substantially higher odds of having sepsis within 3 months after PCDF compared with ACDF (N=3,714, 2.5% vs. 0.7%, adjusted OR 3.56, 95% CI 1.96 to 6.91)[43] while the other two studies (N=3,588) reported similar rates between groups (ACDF, range <0.7% to 1.1% vs. PCDF/posterior laminoplasty, range <0.7% to 1.7%)[45,47]

Surgical site infection was reported by four NRSIs. Three studies (N=22,702)[41,46,47] reported that posterior approaches (PCDF or laminoplasty) were associated with a large increase in the odds of surgical site infection compared with ACDF at 1 to 3 months (frequency range 2.4% to 4.7% vs. 0.8% to 1.0%, OR range 3.1 to 3.7) and the fourth (N=245) found no difference between groups (1% each).[39]

Wound dehiscence was rare across four NRSIs, two of which reported that PCDF was associated with a substantial increase in the odds of experiencing this complication compared with ACDF (N=19,660, frequency range 1.3% to 2.7% vs. 0.1% to 0.5%, range of ORs 5.6 to 10.8)[41,46] and two that found no difference between groups, (1% each, N=245, 1 RCT) [39] and (0% each, N=264, 1 RCT).[42]

Dural tear/durotomy occurred more often with ACDF versus PCDF in one study (N=627, 9.4% vs. 3.2%, RR 3.02, 95% CI 1.50 to 6.10)[44] while no events were reported in either group in another study (N=264).[42]

One NRSI found that PCDF was associated with a large increase in the odds of having ***any severe adverse*** event through 3 months compared with ACDF (N=3,714, 13% vs. 6.1%, OR 2.31, 95% CI 1.83 to 2.93).[43]

A variety of ***other serious adverse events*** were reported across five NRSIs (total N=21,813, range 546 to 13,884);[40,43-45,47]; event rates ranged from 0.04% to 4.5% in the ACDF arms and from 0% to 7.7% in the posterior arms (PCDF or laminoplasty) and included kidney injury (4 studies)[43-45,47] cardiac complications (4 studies),[40,44,45,47] transfusion (3 studies),[43-45] respiratory complications (3 studies),[40,44,47] and arterial injury and hardware instrument failure malposition (1 study).[40] Excluding perioperative blood transfusion in one study, which had the highest frequency of events across all these complications (N=627, 4.5% with ACDF vs. 7.7% with a posterior approach),[44] the range across treatment arms was 0% to 3.7% (ACDF) versus 0.06% to 3.6% (posterior approach). There were no cases of myocardial infarction or vocal cord paralysis in one NRSI (N=245).[39]

3.8 Key Question 7. In patients with cervical spondylotic myelopathy due to cervical degenerative disease, what are the comparative effectiveness and harms of cervical laminectomy and fusion compared to cervical laminoplasty in patients?

3.8.1 Key Findings

- Evidence was inadequate to determine the effect of laminectomy versus laminoplasty on neck, shoulder, or arm pain (SOE: Insufficient).
- There was moderate-strength evidence of little difference between laminectomy and fusion versus laminoplasty on neurologic function (SOE: Moderate) and low-strength evidence of no difference between laminectomy and fusion versus laminoplasty on general function (SOE: Low).
- There was moderate-strength evidence of no difference in reoperation rates between laminectomy and fusion compared with laminectomy (SOE: Moderate).
- There was low-strength evidence of fewer complications with laminoplasty compared with laminectomy and fusion (SOE: Low).

3.8.2 Description of Included Studies

Two RCTs (N=46)[48,49] and 6 NRSI (N=15,523)[50-55] compared cervical laminectomy and fusion with cervical laminoplasty (**Appendix C**). The followup duration was 1 year in both of the randomized trials and ranged from 1 year to 5 years in the nonrandomized studies. Trials were conducted in the U.S. and Egypt, with NRSI studies conducted in the U.S. (3 studies), Japan, China, and a multinational setting.

The mean age of participants was 58 years in one trial and not reported in the other (most participants in the second trial ranged from 50 to 59 years); mean ages in the nonrandomized studies ranged from 54 to 64 years. The average proportion of females in the trials was 30% and 58%; the proportion of females in the NRSI studies ranged from 21% to 55%. Race and ethnicity were not reported in any of the studies. One trial enrolled patients with at least 3 levels of spinal cord compression,[48] while the other did not report the number of disease levels.[49] Two nonrandomized studies enrolled patients with 3 or more levels of spinal cord compression,[52,55] whereas the number of disease levels was not specified in the other NRSI studies.

One RCT was rated high risk of bias[48] and the other was rated as moderate risk of bias.[49] All of the observational studies were rated moderate risk of bias (**Appendix D**). The evidence comparing laminectomy and fusion with laminoplasty for neck, shoulder, and arm pain was rated insufficient due to limited and conflicting evidence (**Appendix G**).

3.8.3 Detailed Analysis

3.8.3.1 Fusion

No study reported fusion outcomes.

3.8.3.2 Pain

There was inadequate evidence to determine the benefits and harms of laminectomy and fusion compared with laminoplasty on neck, shoulder, or arm pain (SOE: Insufficient).

One RCT (N=30) found a moderate benefit in neck pain with laminectomy and fusion compared with laminoplasty at one year (MD -1.33, $p<0.05$) but no difference in limb pain (MD 0.4, $p>0.05$).[48] The other randomized trial (N=16) reported improvement in neck and arm pain from baseline only in patients who underwent laminoplasty (surgical approaches not directly compared, numeric values not reported, $p<0.05$, both outcomes).

Among the nonrandomized studies assessing neck[50,52] or shoulder[50] pain, two (N=148) reported no differences in VAS scores between laminectomy and fusion and laminoplasty at 1 or 3 years.[50,52] Another observational study (N=121) reported no differences in improved pain (74% vs. 60%; p=0.141) for posterior laminectomy and fusion versus laminoplasty.[55]

3.8.3.3 Function

3.8.3.3.1 Neurologic Function

There was moderate-strength evidence of no difference between laminectomy and fusion versus laminoplasty on neurologic function (SOE: Moderate).

Two head-to-head RCTs (N=46) assessed neurologic function with the mJOA and the Nurick Classification Scale for Spinal Cord Compression (i.e., Nurick's grade 0 to 5) at 1 year post-operative.[48,49] Pooled analysis of the two trials found no difference in function between cervical laminectomy and fusion versus laminoplasty using the mJOA (N=46, MD -0.03, 95% CI, -0.68 to 0.74, I^2=76%).[48,49] One trial reported no significant difference between laminectomy and fusion compared with laminoplasty in Nurick grade (1.40 vs. 1.67; p=0.23),[48] while the other trial reported a significant pre-post difference for laminoplasty only (numeric values not reported; $p<0.05$).[49]

Four nonrandomized studies reported neurologic function using the mJOA or JOA score; three reported no difference between laminectomy and fusion versus laminoplasty[50,52,55] and one reported a significant benefit of laminoplasty over laminectomy and fusion (mean mJOA at 2 years: 3.49, 95% CI 2.84 to 4.13 vs. 2.39, 95% CI 1.91 to 2.86; p=0.0069).[51] However, this study reported no significant difference in Nurick's grade at 2 years (1.57, 95% CI 1.23 to 1.90 vs. 1.18, 95% CI 0.92 to 1.44; p=0.077).

3.8.3.3.2 General Function

There low-strength evidence of little difference between laminectomy and fusion versus laminoplasty on general function (SOE: Low).

Neck disability scores on the NDI were not different between laminectomy and fusion versus laminoplasty (1 RCT, N=30, MD 3.86, p=0.2)[48] and only improved with laminoplasty in the other trial (N=16, surgical approaches not directly compared, numeric values not reported, p=0.05).[49] The same trial (N=16) reported improvement from baseline on the SF-36 with laminoplasty only (numeric values not reported, $p<0.05$).[49]

[50,52]Two NRSI reported no differences on the NDI,[50,51] and three reported no differences between surgical approaches in SF-12 or SF-36 PCS or MCS scores.[50-52] Another observational study reported no differences in improved gait (71% vs. 68%; p=0.674) as assessed on a 5-point NRS.[55]

3.8.3.4 Quality of Life

No study reported quality of life outcomes.

3.8.3.5 Harms

There was moderate-strength evidence of no difference between laminectomy and fusion compared with laminectomy in reoperation rates (SOE: Moderate) and low-strength evidence of fewer complication overall with laminoplasty compared with laminectomy and fusion (SOE: Low).

Both trials reported no significant differences in harms, though event rates were low.[48,49] Likewise, four NRSI studies (N=582) found no differences in infection, device failure, or reoperation rates.[50-52,55] A large database study (PearlDiver Mariner Database, N=11,860, unsure of matched sample size)[53] reported similar revision rates for laminoplasty and laminectomy with fusion (5.63% vs. 5.90%, p=0.62) at 1 year but fewer surgical site infections (matched OR 0.60; p=0.002), wound complications (matched OR 0.67, p=0.002) and dysphagia (matched OR 0.77; p=0.01) with laminoplasty compared with laminectomy and fusion.[53] Also reported in this study were reduce rates of spinal cord injury (matched OR 0.6, p=0.02), limb paralysis (matched OR 0.67, p<0.001), respiratory failure (matched OR 0.74, p=0.01), renal failure (matched OR 0.84, p=0.04), and sepsis (matched OR 0.85, p=0.04) with laminoplasty versus laminectomy and fusion. No complication was reported more likely with laminoplasty. An earlier propensity-matched analysis of patients from this same database (N=928) found lower revision rates at 1 year with laminoplasty versus laminectomy and fusion (2.4% vs. 7.1%; p<0.001).[54] The dissimilar findings may be due a larger sample size (this is an assumption as the matched sample size was not reported in the later study) to changes in surgical methods and/or skill of the surgeon over time. Two additional NRSI studies reported no differences in dysphagia between groups.[51,55]

3.9 Key Question 8. In patients with cervical spondylotic radiculopathy or myelopathy at one or two levels, what are the comparative effectiveness and harms of cervical arthroplasty compared to anterior cervical discectomy and fusion?

3.9.1 Key Findings

- In participants receiving single-level interventions:
  - There was moderate-strength evidence of no difference between cervical artificial disc replacement (C-ADR) and ACDF in likelihood of success (response) for any pain or function measure at short, intermediate, and long term (SOE: Moderate).
  - There was also moderate-strength evidence of no differences between C-ADR and ACDF in pain or function at short, intermediate, or long term: neck or arm pain, neurologic status or general function (SOE: Moderate).
  - There was high-strength evidence that C-ADR was associated with substantially lower likelihood of reoperation at the index level versus ACDF (SOE: High).
  - There was low-strength evidence that C-ADR was associated with slightly lower likelihood of any serious AE at short term versus ACDF, but there were no differences at times >24 months and serious AEs were variably defined (SOE: Low for all times).
  - There was low-strength evidence of no differences in neurological events or deficits between C-ADR and ACDF at short, intermediate, or long term (SOE: Low).
  - There was inadequate evidence on the likelihood of mortality between C-ADR and ACDF (SOE: Insufficient).

- In participants receiving 2-level interventions:
  - There was moderate-strength evidence of no differences between C-ADR and ACDF on pain (neck or arm), neurologic function and general function at short, intermediate, and long term (SOE: Moderate).
  - Reoperation at the index level was substantially less likely with C-ADR at all times reported (24 to >60 months) (SOE: Low).
  - C-ADR was associated with slightly lower likelihood of serious adverse events compared with ACDF at 24 months, but there was no difference between procedures at 120 months for WHO Grade 3 or 4 adverse events (SOE: Low).
    Evidence for neurological deficits or events and for mortality was inadequate to draw conclusions (SOE: Insufficient).

- In participants receiving 1-, 2- or 3-level interventions
  - There was no difference between C-ADR and ACDF in VAS neck pain scores at intermediate term (SOE: Low).
  - Evidence was inadequate to draw conclusions for neurologic and general function and harms (SOE: Insufficient).

3.9.2 Description of Included Studies

Twenty-two RCTs in 45 publications (N=4,120) compared cervical arthroplasty (C-ADR) to anterior cervical discectomy and fusion (ACDF) (**Appendix C**).[56-100] The average followup duration was 56 months (range 6 to 108 months). Eight trials each were conducted in the U.S.[63,70,73,74,84,85,91,96] and in China;[59-61,77,89,97-99] two trials in Germany;[87,88] and one trial each in India,[72] the Netherlands,[101] Spain,[62] and Turkey.[80]

The average study mean age of participants was 45 years (range 37 to 50 years); the average proportion of females in studies was 47% (range 20% to 63%). Five trials reported race, four enrolling mostly White participants (range 89 to 93%)[70,74,91,96] and the other enrolling Han (Chinese) participants.[61] One trial reported ethnicity, enrolling mostly non-Hispanic participants (94%).[63]

Studies enrolled participants with clinical and/or radiological evidence of cervical radiculopathy and/or myelopathy, although only three trials reported baseline values.[62,72,87] Participants had 1-level disease in 15 trials (N=3,036),[59,73,74,77,80,84,85,87-89,91,96,98,99,101] 2-level disease in four trials (N=872),[61,63,70,97] and mixed-level (1, 2 or 3) disease in three trials (N=196).[60,62,72] Of the single-level trials, six (in 23 publications) were U.S. Food and Drug Administration (FDA) Investigational Device Exemption (IDE) trials[56-58,65,66,68,73-76,78,79,82-85,90,91,93-96,100] and of the 2-level trials, two (in 9 publications) were IDE trials.[63,64,69-71,78,81,92,93]

Six trials were rated low risk of bias,[63,74,77,84,85,91] six trials were rated high risk of bias,[59,62,80,88,89,99] and the remainder were rated moderate risk of bias[60,61,70,72,73,87,96-98,101] (**Appendix D**). Methodological limitations included unclear randomization techniques, unclear blinding, and high attrition.

Two prospective, multicenter NRSIs (N=349 and N=352) of recently completed FDA IDE trials compared newer C-ADR devices (M6-C and Simplify discs) with historic ACDF controls (**Appendix C**).[102,103] Propensity score matching was done to facilitate baseline comparability between groups. Follow-up was 24 months in both studies. One study enrolled participants with clinical and radiological evidence of cervical radiculopathy with or without myelopathy at 1-level[103] and the other study enrolled participants with cervical radiculopathy and/or myelopathy at 2-levels.[102] The study mean ages of participants were 45 years and 48 years and the proportion of females were 50% and 52%. Race/ethnicity was not reported by either study. The study mean BMIs were 27.5 and 28.9. Both studies were conducted in the United States and were rated moderate risk of bias (**Appendix D**).

Six non-IDE NRSIs were included for the evaluation of harms only and included five large database/registry studies,[104-108] one a post-hoc analysis of an FDA IDE trial[109] (**Appendix C**). Sample sizes ranged from 342 to 143,060 (total N=204,505). The average study mean age of patients was 50 years (range 46 to 54 years) and the proportion of females was 51% (range 50% to 52%). Across three studies most patients were White (82%; range 81% to 85%); one study reported 94% of patients were non-Hispanic[109] and two studies did not report race/ethnicity.[107,108] Two studies[105,109] enrolled patients with radiculopathy and/or myelopathy; one study[104] specifically excluded patients with myelopathy and the remaining three studies[106-108] only stated that patients had cervical degenerative disease. Follow-up ranged from 30 days to 84 months. One study took place in Germany,[108] and all others in the United States. Two studies were rated moderate risk of bias[105,109] and four high risk of bias[104,106-108] (**Appendix D**).

For the FDA IDE trials, an attempt was made to reconcile conflicting information among multiple reports presenting the same data and when necessary, we used the data from the FDA Summary of Safety and Effectiveness Data (SSED): 1-level[110-116] and 2-level indications.[117-119]

For measures of success, we focused on the FDA required definition and reported alternative definitions as applicable. Only FDA approved devices are included for this key question.

In the results below for benefits, we report outcomes according to the following timeframes: short term (<12 months), intermediate term (12 to 60 months) and long term (>60 months).

Evidence was insufficient for mortality (all levels), neurologic deficit/events (2-levels and mixed 1-, 2- or 3-levels), and neurologic function, general function, reoperation and serious AEs (mixed 1-, 2- or 3-levels) based on a combination of two or more of the following: high risk of bias, inconsistent findings, and lack of precision (**Appendix G**).

3.9.3 Detailed Analysis

3.9.3.1 Single-level C-ADR versus ACDF

Fifteen trials (N=3,036) (in 33 publications) compared single-level C-ADR and ACDF, including six FDA IDE trials (in 23 publications)[56-58,65,66,68,73-76,78,79,82-85,90,91,93-96,100] and nine non-IDE trials (in 10 publications),[59,77,80,87-89,98,99,101] as did one FDA IDE NRSI.[103] Six additional NRSIs compared harms for single-level C-ADR and ACDF.[104-109]

3.9.3.1.1 Fusion

Seven RCTs (across 15 publications) (N=2,382) that compared single-level C-ADR and ACDF reported fusion success in their ACDF arms.[57,58,66,73-76,84,85,90-93,96,99] One trial (N=56) reported short-term fusion success in 89.3 percent of participants,[99] seven RCTs (N=853) reported intermediate-term fusion success in 93.9 percent (range 89.1% to 98.2%) of participants[57,66,73,76,90,96,99] and two RCTs (N=181) reported long-term fusion success in 96.5 percent (range 95.5% to 96.9%) of participants.[58,93] One RCT reported successful fusion in the C-ADR arm as well, but this may be attributed to participant crossover after initial randomization.[90,91]

3.9.3.1.2 Pain

3.9.3.1.2.1 Neck pain

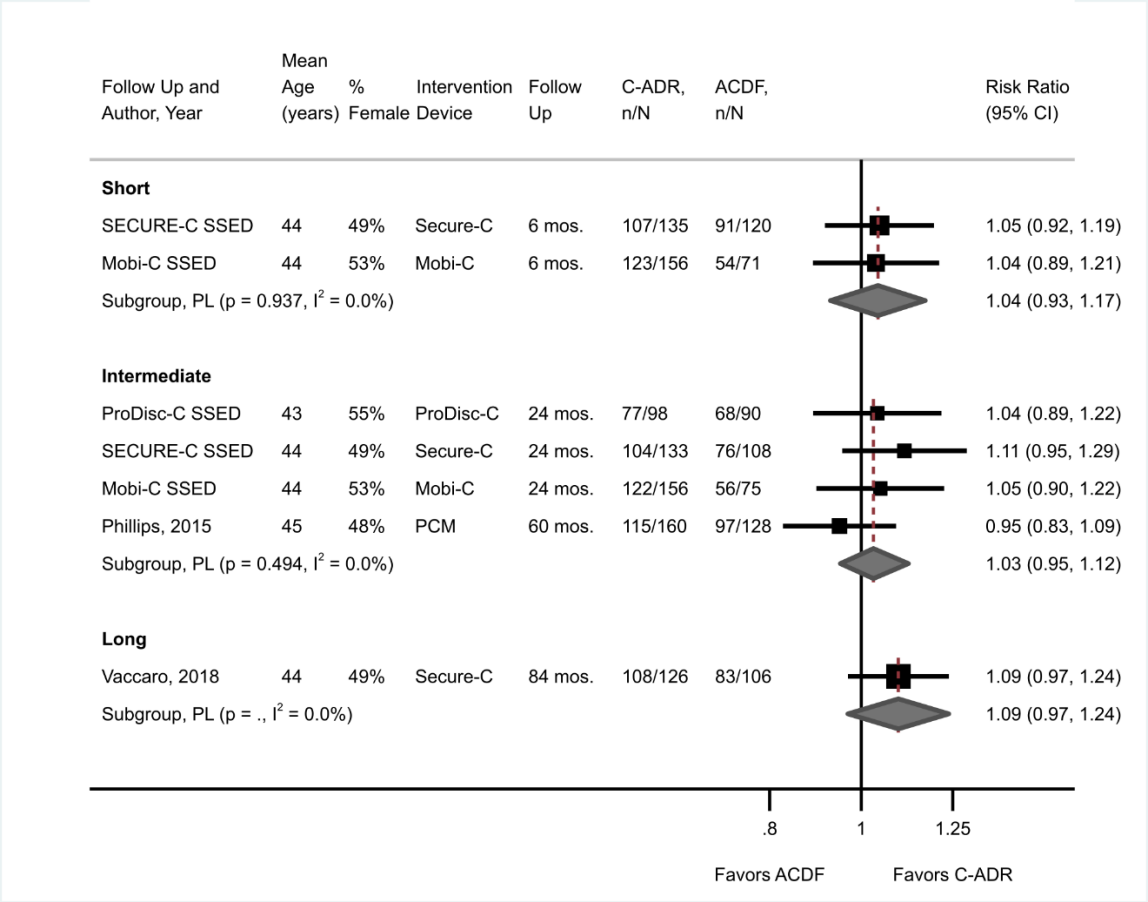
There was moderate-strength evidence of no differences between C-ADR and ACDF in neck pain or likelihood of success (response) for neck pain at short, intermediate, and long-term (SOE: Moderate).

Four RCTs (N=1,230) (in 5 publications)[90,95,110,114,115] that compared single level C-ADR versus ACDF reported neck pain success (response) defined as postoperative ≥20-point improvement on VAS. There were no differences in likelihood of neck pain success between C-ADR and ACDF at short term (2 RCTs, N=482, 79% vs. 75.0%, RR 1.04, 95% CI 0.93 to 1.17, I^2=0%),[110,115] intermediate term (4 RCTs, N=948, 76.4% vs. 74.1%, RR 1.03, 95% CI 0.95 to 1.12, I^2=0%)[90,110,114,115] or long term (1 RCT, N=232, 85.7% vs. 78.3%, 1.09, 95% CI 0.97 to 1.24)[95] (**Figure 7**). In one prospective NRSI IDE study using propensity-matched historical controls, more C-ADR participants had ≥20-point improvement on VAS neck pain versus ACDF at 24 months (N=301, 91.2% vs. 77.9%, p=0.013).[116]

One of the above trials reported neck pain success at 84 months using an alternative definition, a ≥10-point improvement on VAS, and was not included in the meta-analysis at long term; there was no difference between C-ADR and ACDF using this criterion (N=191, 87.5% vs. 83.3%, RR 1.05, 95% CI 0.93 to 1.20).[93]

3. Results

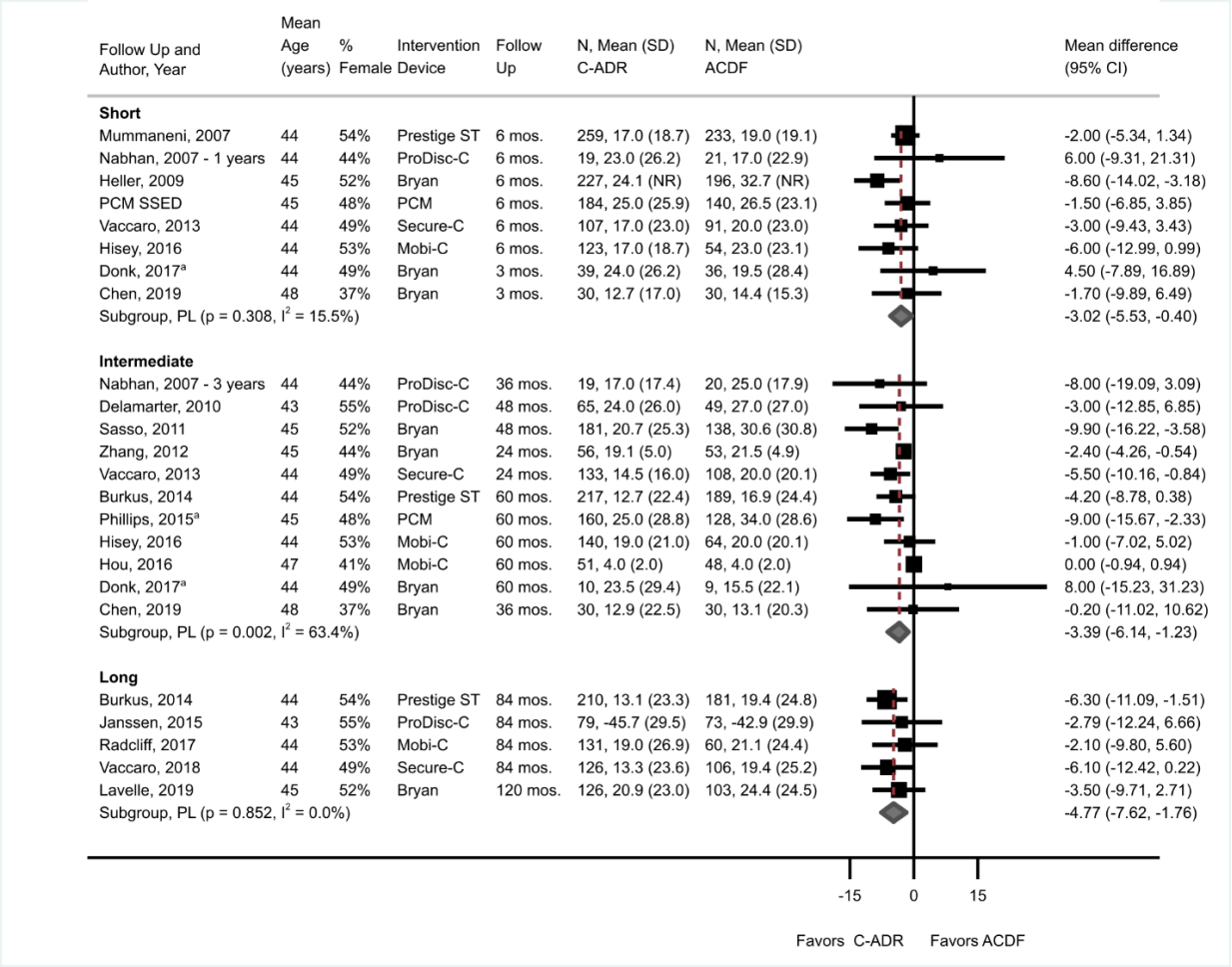
Figure 7. Neck pain success (≥20-point improvement on VAS): Comparison of C-ADR with ACDF (1-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PL = profile likelihood; SSED = Summary of Safety and Effectiveness Data (FDA); VAS = visual analog scale.

Eleven RCTs (N=2,696) (in 19 publications)[58,59,65,67,73,76,77,79,82,84,86,87,90,93-96,98,112] contributed to evaluation of mean differences in neck pain scores at various times. There were no differences between C-ADR and ACDF in VAS neck pain scores (0-100 scale) as estimates were below the threshold for a small effect at short term (8 RCTs, N=1,789, MD -3.02, 95% CI -5.53 to 0.40, I^2=15.5%),[59,67,73,76,84,87,96,112] intermediate term (11 RCTs, N=1,898, MD -3.39, 95% CI -6.14 to -1.23, I^2=63.4%)[58,59,65,67,76,77,86,90,94,96,98] and long-term (5 RCTs, N=1,195, MD -4.77, 95% CI -7.63 to -1.76, I^2=0%)[58,79,82,93,95] (**Figure 8**). Exclusion of one, small (N=60) trial rated high risk of bias[59] did not substantially change effect estimates but did slightly increase heterogeneity in the short term (7 RCTs, N=1,729, MD -3.11, 95 % CI -5.92 to -0.15, I^2=26.6%)[67,73,76,84,87,96,112] and intermediate term (10 RCTs, N=1,838, MD -3.55, 95% CI -6.48 to -1.30, I^2=67.1%).[58,65,67,76,77,86,90,94,96,98] Exclusion of one trial[67] that did not specify if neck or arm pain was evaluated also did not substantially change effect estimates at short term (7 RCTs, N=1,714, MD -3.24, 95% CI -5.95 to -0.77, I^2=12.2%)[59,73,76,84,87,96,112] or intermediate term (10 RCTs, N=1,879, MD -3.51, 95% CI -6.35 to -1.33, I^2=66.4%).[58,59,65,76,77,86,90,94,96,98] Although funnel plot analysis and Egger's test (p=0.035) may suggest publication/small study bias for neck pain scores at intermediate term, most trials found no effect leading to less concern regarding publication bias (**Appendix F, Figure 1**).

Figure 8. Neck pain VAS scores (0-100 scale): Comparison of C-ADR with ACDF (1-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PL = profile likelihood; SD = standard deviation; SSED = Summary of Safety and Effectiveness Data (FDA); VAS = visual analog scale.

[a] Scores estimated from graphs in article.

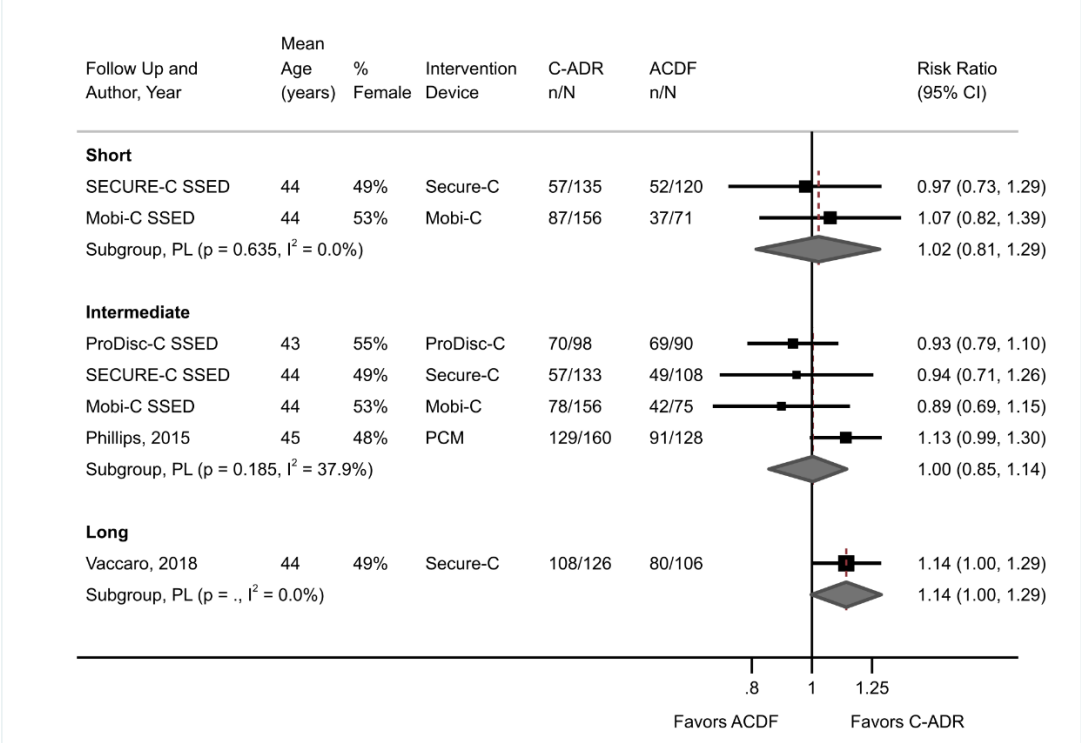
3.9.3.1.2.2 Arm pain

There was moderate-strength evidence of no differences between C-ADR and ACDF in arm pain or likelihood of success (response) for arm pain at short, intermediate, and long-term (SOE: Moderate).

Four RCTs (N=1,148) (in 5 publications)[90,95,110,114,115] that compared C-ADR with ACDF for single level disease reported ***arm pain success (response)*** defined as postoperative ≥20-point improvement on VAS (0–100). Some studies reported arm pain success in both arms. Conservative estimates, using the lower risk ratio for studies reporting VAS for both arms, revealed no difference in likelihood of arm pain success between C-ADR and ACDF at short-term (2 RCTs, N=482, 49.5% vs. 46.6%, RR 1.02, 95% CI 0.81 to 1.29, I^2=0%),[110,115] intermediate (4 RCTs, N=948, 61.1% vs. 62.6%, RR 1.0, 95% CI 0.85 to 1.14, I^2=37.9%)[90,110,114,115] or long-term (1 RCT, N=232, 85.7% vs. 75.5%, RR 1.14, 95% CI 1.0 to 1.29, I^2=0%)[95] (**Figure 9**). Estimates based on higher risk ratios for studies reporting VAS for both arms were similar and led to the same conclusion of no difference between C-ADR and

ACDF for all time points. In one prospective NRSI IDE study using propensity-matched historical controls, more C-ADR participants experience ≥20-point improvement on VAS arm pain (worst side) versus ACDF at 24 months (N=301, 90.5% vs. 79.9%, p=0.001).[116]

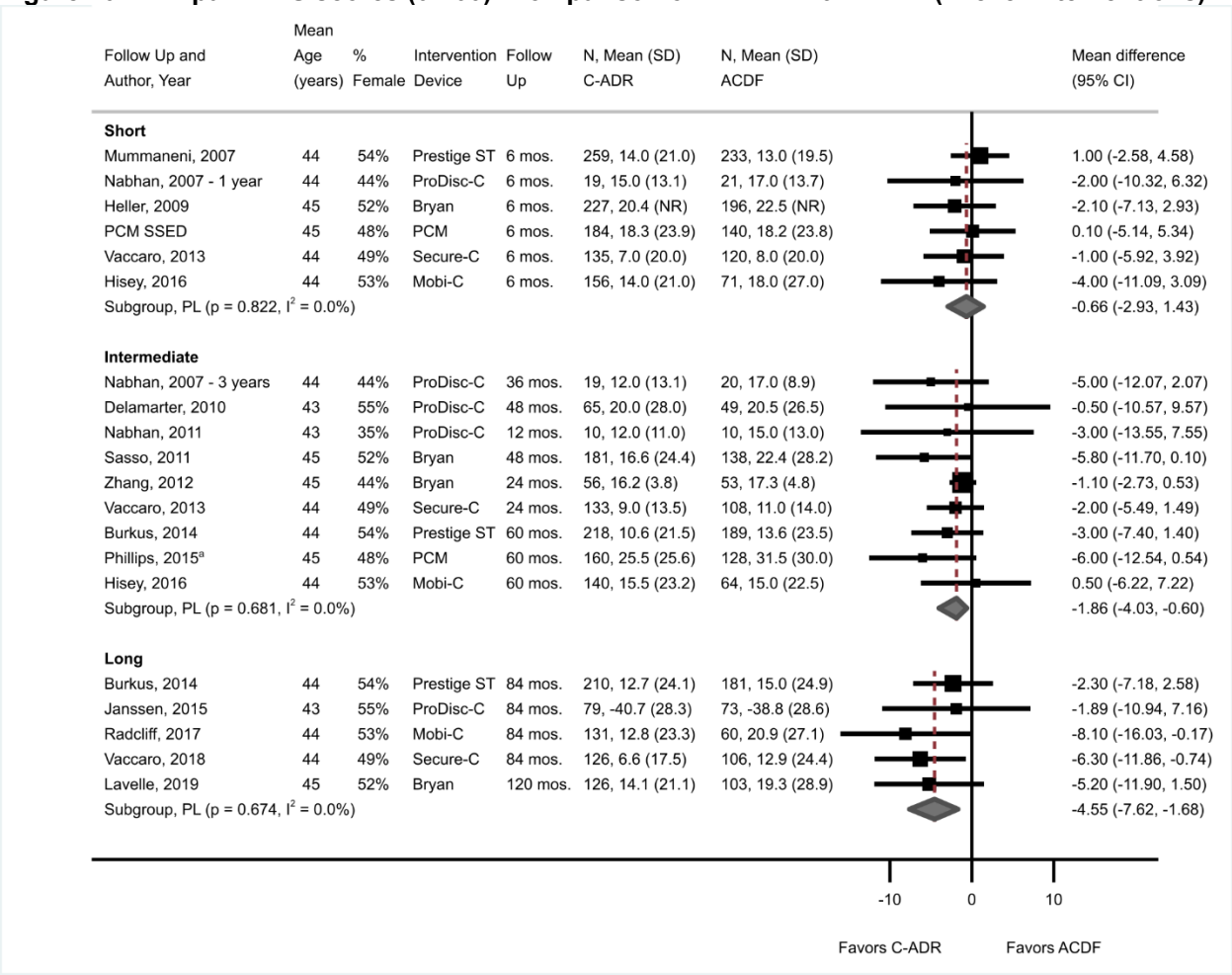
Figure 9. Arm pain success (≥20-point improvement on VAS): Comparison of C-ADR with ACDF (1-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PL = profile likelihood; SSED = Summary of Safety and Effectiveness Data (FDA); VAS = visual analog scale.

Nine RCTs (N=2,460) (in 17 publications)[58,65,73,76,79,82,84,86-88,90,93-96,98,112] assessed arm pain at various times. Three publications reported pain scores for both arms. Using a conservative estimate with the smaller effect estimate of the two arms, there was no difference between C-ADR and ACDF in VAS arm pain scores (0-100 scale) short term (6 RCTs, N=1,761, MD -0.66, 95% CI -2.93 to 1.43, I^2=0%),[73,76,84,87,96,112] intermediate term (9 RCTs, N=1,741, MD -1.86, 95% CI -4.03 to -0.56, I^2=0%)[58,65,76,86,88,90,94,96,98] or long-term (5 RCTs, N=1,195, MD -4.55, 95% CI -7.62 to -1.68, I^2=0%)[58,79,82,93,95] (**Figure 10**). Exclusion of one small (N=20) trial rated high risk of bias[88] did not impact the effect size. Using the larger effect estimate when both arms were measured, slightly increased the estimate at short term but not the conclusion of no difference between treatments (MD -1.11, 95% CI -3.56 to 1.02); estimates at intermediate and long term were similar to the conservative estimates.

Figure 10. Arm pain VAS scores (0-100): Comparison of C-ADR with ACDF (1-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; PL = profile likelihood; SD = standard deviation; SSED = Summary of Safety and Effectiveness Data (FDA); VAS = visual analog scale.

[a] Scores estimated from graphs in article.

3.9.3.1.3 Function

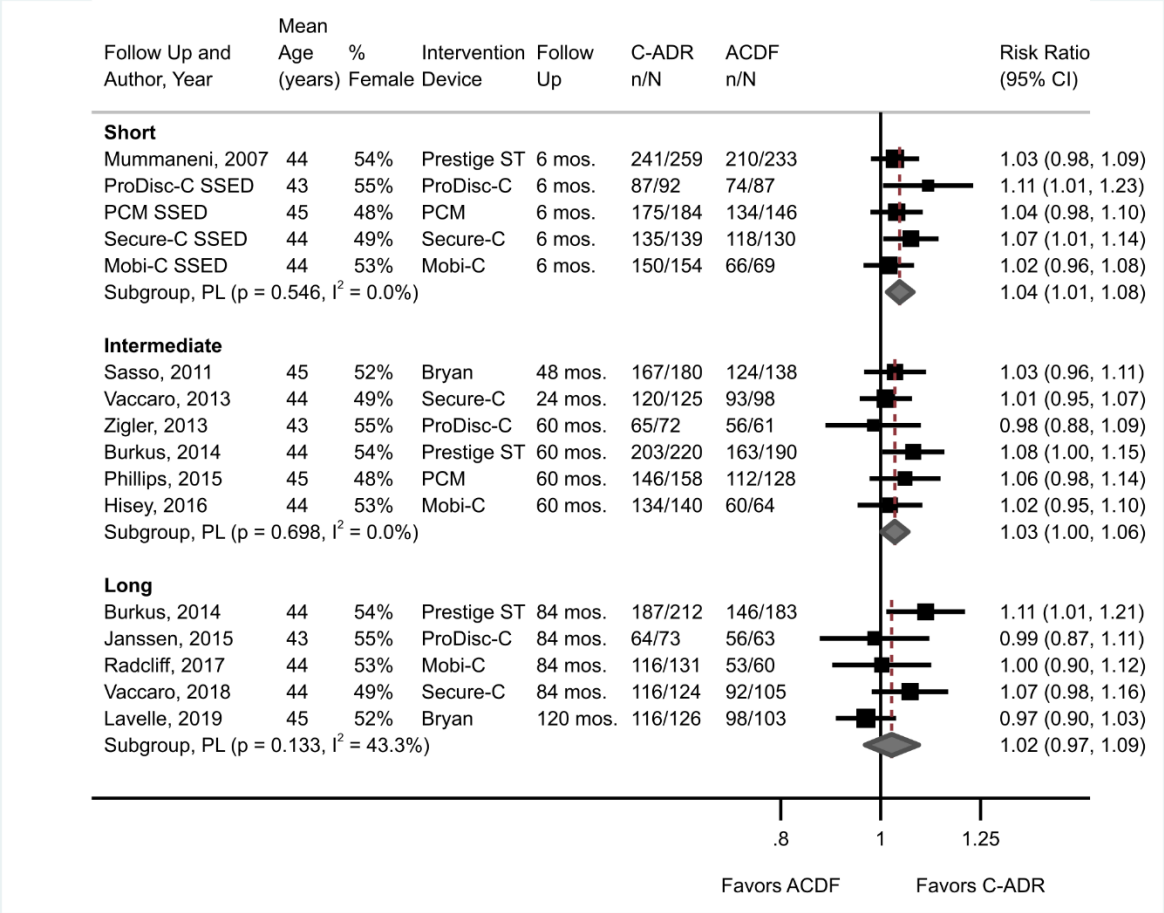
3.9.3.1.3.1 Neurologic Function

There was moderate-strength evidence of no differences between C-ADR and ACDF in neurologic function at short, intermediate, and long-term (SOE: Moderate).

Six RCTs (N=2,271) (in 15 publications)[58,76,79,82,84,90,93-96,100,110,112,114,115] that compared single-level C-ADR and ACDF reported ***neurologic success (response)*** defined as maintenance or improvement (compared with preoperative status) in all three of the following areas: motor function, sensory function and deep tendon reflexes. There were no differences between C-ADR and ACDF in the likelihood of neurological success short-term (5 RCTs, N=1,493, 95.2% vs. 90.5%, RR 1.04, 95% CI 1.01 to 1.08, I^2=0%),[84,110,112,114,115] intermediate term (6 RCTs, N=1,574, 93.3% vs. 89.5%, RR 1.03, 95% CI 1.00 to 1.06, I^2=0%)[58,76,90,94,96,100] or long term (5 RCTs, N=1180, 89.9% vs. 86.6%, RR 1.02, 95% CI 0.97 to 1.09, I^2=43.3%)[58,79,82,93,95] (**Figure 11**). One prospective NRSI IDE study that used propensity matched ACDF historical controls

reported neurological success, defined as maintenance or improvement compared with baseline, was similar for C-ADR and ACDF at 24 months (N=314, 99.3% vs. 98.8%).[116]

Figure 11. Neurological success: Comparison of C-ADR with ACDF (1-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PL = profile likelihood; SSED = Summary of Safety and Effectiveness Data (FDA); VAS = visual analog scale.

Four RCTs (N=354), three rated high risk of bias[61,89,99] and one low risk of bias,[77] reported ***JOA scores*** (0-17). There was no differences between C-ADR and ACDF in pooled analysis at intermediate term (4 RCTs, N=354, MD 0.60, 95% CI -0.007 to 0.97, I^2=1.9%) or in one short-term trial rated high risk of bias (1 RCT, N=60, MD 0.25, 95% CI -0.25 to 0.75).[61]

One trial reported the proportion of participants who had the same or an improved ***Nurick grade*** at 60 months compared with baseline; there were no differences (i.e., point estimate below the threshold for a small effect) between C-ADR and ACDF (N=285, 99.4% vs. 96.9%, RR 1.03, 95% CI 0.99 to 1.06).[91]

3.9.3.1.3.2 General Function

There was moderate-strength evidence of no differences between C-ADR and ACDF in general function at short, intermediate, and long-term (SOE: Moderate).

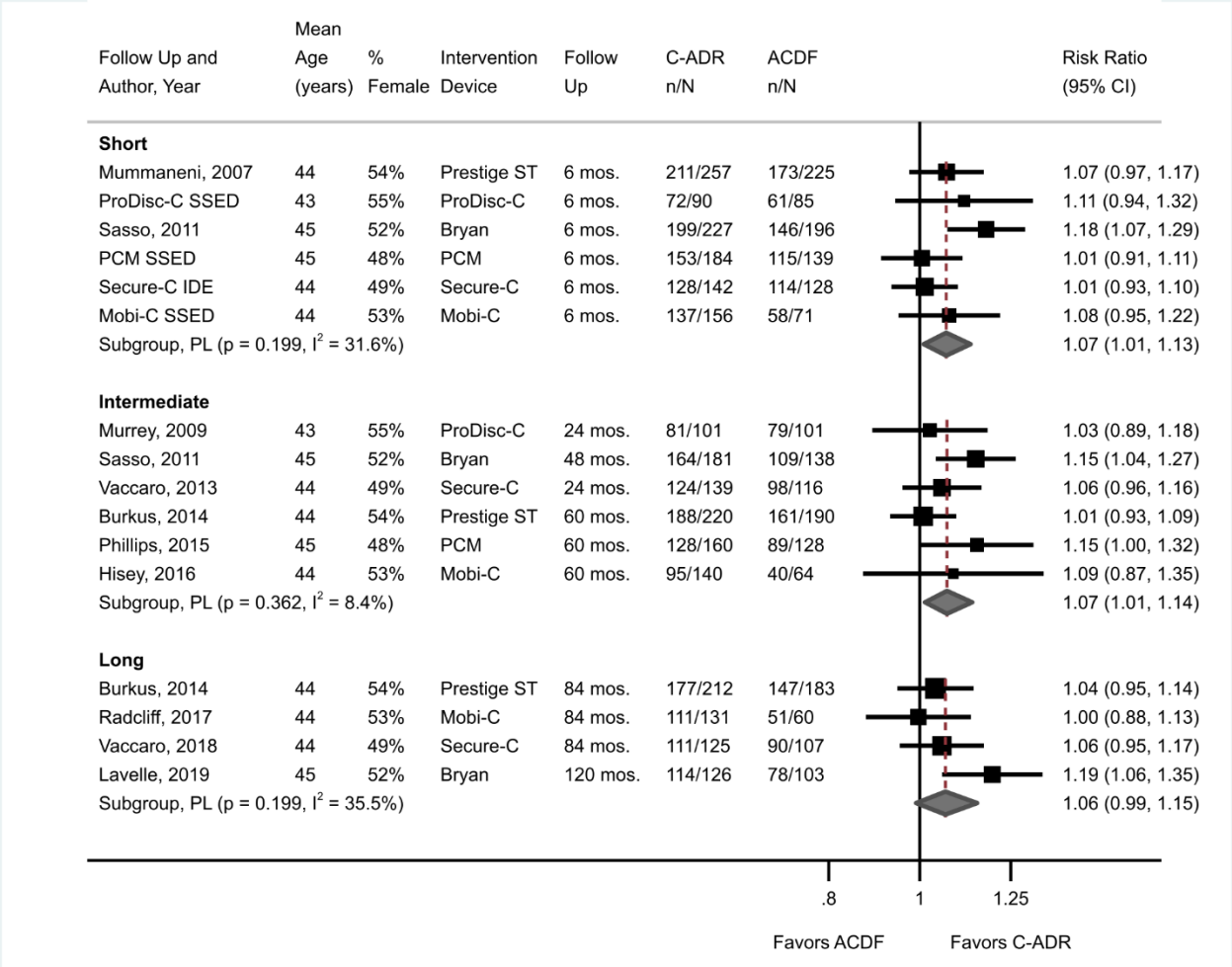
3.9.3.1.3.2.1 NDI

Six RCTs (N=2,271) (in 14 publications)[58,76,82,84,85,90,93-96,110,112,114,115] that compared C-ADR with ACDF for single-level disease reported NDI success (response) defined as postoperative

3. Results

NDI score improvement of ≥15 points from the baseline score (FDA definition). There were no differences between C-ADR and ACDF in the likelihood of NDI success short term (6 RCTs, N=1,900, 85.2% vs. 79.0%, RR 1.07, 95% CI 1.01 to 1.13, I^2=0%),[84,94,110,112,114,115] intermediate term (6 RCTs, N=1,678, 82.9% vs. 78.2%, RR 1.07, 95% CI 1.01 to 1.14, I^2=8.4%)[58,76,85,90,94,96] or long term (4 RCTs, N=1,047, 86.4% vs. 80.8%, RR 1.06, 95% CI 0.99 to 1.15, I^2=35.5%)[58,82,93,95] (**Figure 12**). In one prospective NRSI IDE study that used propensity-matched historical controls, there was no difference in NDI success (≥15-point NDI improvement) following C-ADR versus ACDF at 24 months (N=301, 90.5% vs. 85.1%, p=0.372).[116]

Figure 12. NDI success (≥15-point improvement): Comparison of C-ADR with ACDF (1-level interventions)

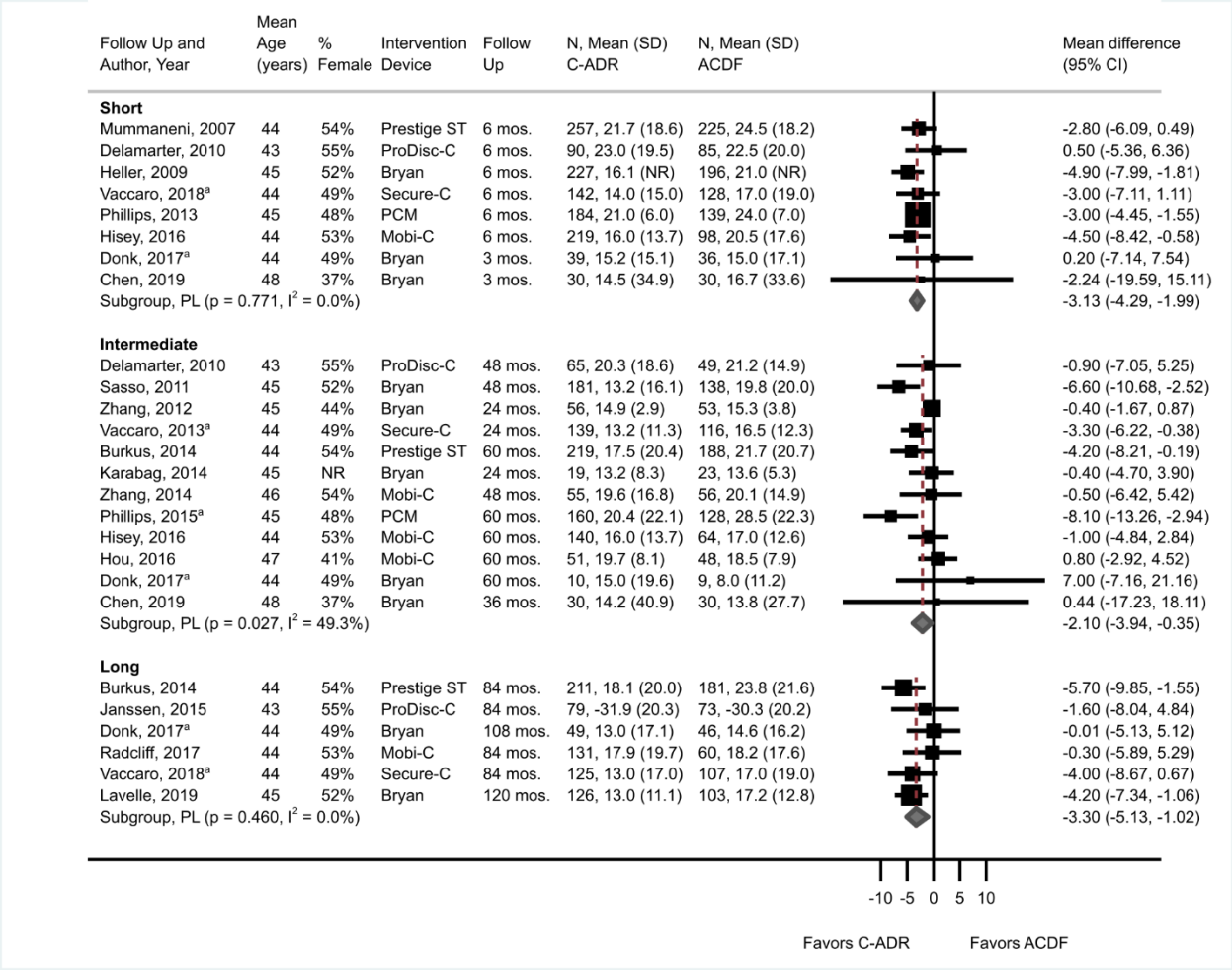


ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; NDI = Neck Disability Index; PL = profile likelihood; SSED = Summary of Safety and Effectiveness Data (FDA).

Twelve RCTs (N=2,800) (in 19 publications)[58,59,65,67,73,76,77,79,80,82,84,90,91,93-96,98,99] that compared C-ADR with ACDF reported NDI scores (0-100 scale). There were no differences between C-ADR and ACDF in NDI scores as estimates were below the threshold for a small effect at short term (8 RCTs, N=2,125, MD -3.13, 95% CI -4.29 to -1.99, I^2=0%),[59,65,67,73,76,84,91,95] intermediate term (12 RCTs, N=2,027, MD -2.10, 95% CI -3.94 to -0.35, I^2=49.3%)[58,59,65,67,76,77,80,90,94,96,98,99] or long-term (6 RCTs, N=1,291, MD -3.30 95% CI -

5.13 to 1.02, I^2=0%)[58,67,79,82,93,95] (**Figure 13**). Exclusion of trials rated high risk of bias[59,80,99] had no impact on effect estimates or statistical heterogeneity in the short-term (7 RCTs, N=2,065, MD -3.14, 95% CI -4.30 to -1.99, I^2=0%)[65,67,73,76,84,91,95] and slightly increased effect size and increased heterogeneity at intermediate term (9 RCTs, N=1,814, MD -2.45, 95% CI -4.70 to -0.35, I^2=62.5%).[58,65,67,76,77,90,94,96,98] Exclusion of a trial rated moderate risk of bias[67] with unclear sample sizes resulted in a small increase in effect size long term (5 RCT, N=1,288, MD -3.78, 95% CI -5.74 to -1.54).[58,79,82,93,95] There was no indication of publication/small study bias for NDI scores at intermediate term based on funnel plot analysis (Egger's test, p=0.416) (**Appendix F, Figure 2**).

Figure 13. NDI scores (0-100): Comparison of C-ADR with ACDF (1-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; NDI = Neck Disability Index; PL = profile likelihood; SD = standard deviation.

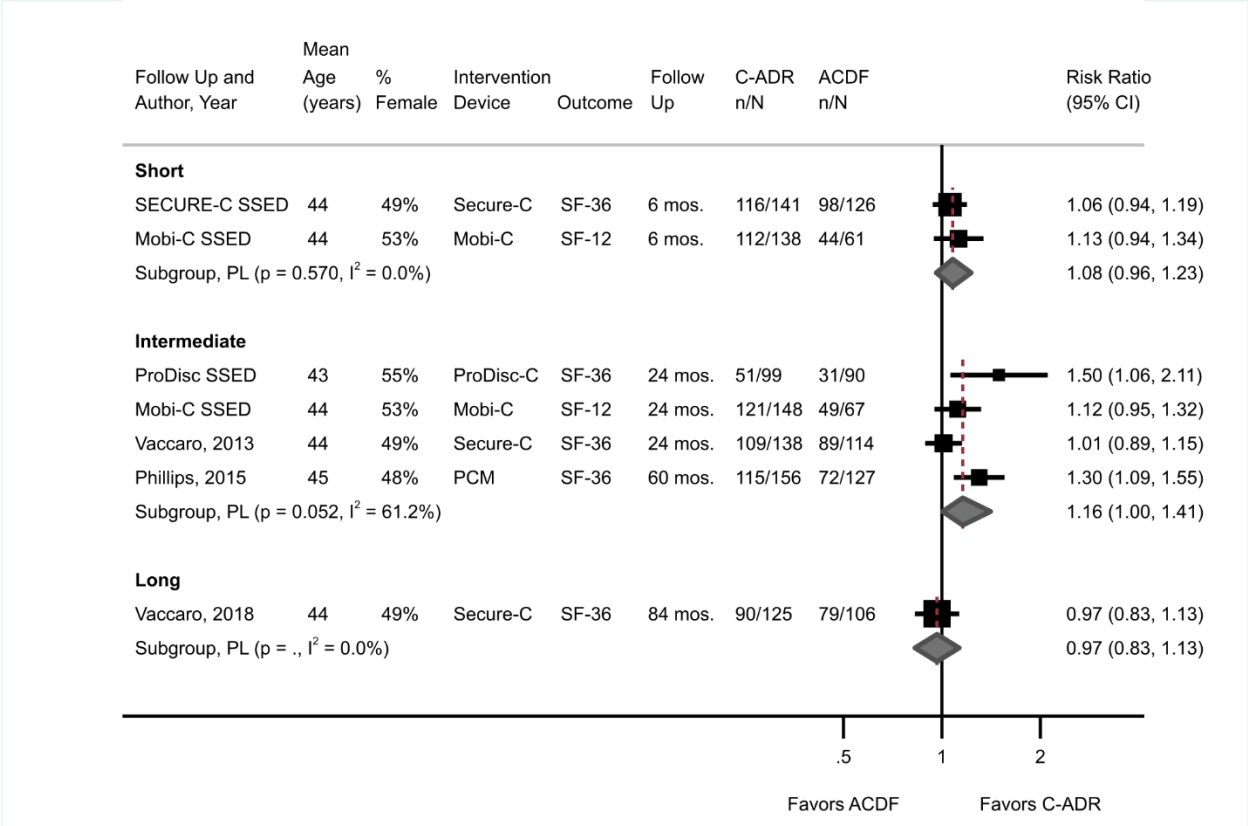
[a] Scores estimated from graphs in article.

3.9.3.1.3.2.2 SF-36 PCS and MCS

Four RCTs (N=1,148) (in 6 publications)[90,95,96,110,114,115] that compared C-ADR with ACDF for single-level disease reported SF-36/12 PCS and MCS (0-100 scale). Success for these component scores was defined as postoperative score improvement of ≥15 points from baseline scores. The likelihood of PCS success was similar for C-ADR and ACDF short term (2 RCTs,

N=466, 81.7% vs. 75.9%, RR 1.08, 95% CI 0.96 to 1.23, I^2=0%),[110,115] intermediate term (4 RCTs, N=939, RR 1.16, 95% CI 1.00 to 1.41, I^2=61.2%)[90,96,110,114] and long term (1 RCT, N=231, 72.0% vs. 74.5%, 0.97, 95% CI 0.83 to 1.13)[95] (**Figure 14**). Exclusion of one outlier trial[114] at intermediate term resulted in a slightly attenuated effect estimate but did not reduce heterogeneity or change the conclusion (3 RCTs, N=750, RR 1.12, 95% CI 0.96 to 1.34, I^2=59.8%).[90,96,110] In one prospective NRSI IDE study using propensity-matched historical controls, more C-ADR participants maintained or improved PCS score versus ACDF at 24 months (N=301, 97.3% vs. 89.2%, p=0.023).[116] The likelihood of MCS success was also similar for C-ADR and ACDF at all time points: short term (2 RCTs, N=466, 49.1% vs. 42.8%, RR 1.13, 95% CI 0.86 to 1.50, I^2=0%),[110,115] intermediate term (4 RCTs, N=939, 47.3% vs. 48%, RR 0.97, 95% CI 0.80 to 1.16, I^2=27.5%)[90,96,110,114] and long term (1 RCT, N=231, 47.2% vs. 43.4%, RR 1.09, 95% CI 0.83 to 1.45)[95] (**Figure 15**). In the prospective NRSI IDE study, there was no difference in MCS maintenance or improvement between procedures at 24 months (N=301, 77.6% vs. 77.0%).[116]

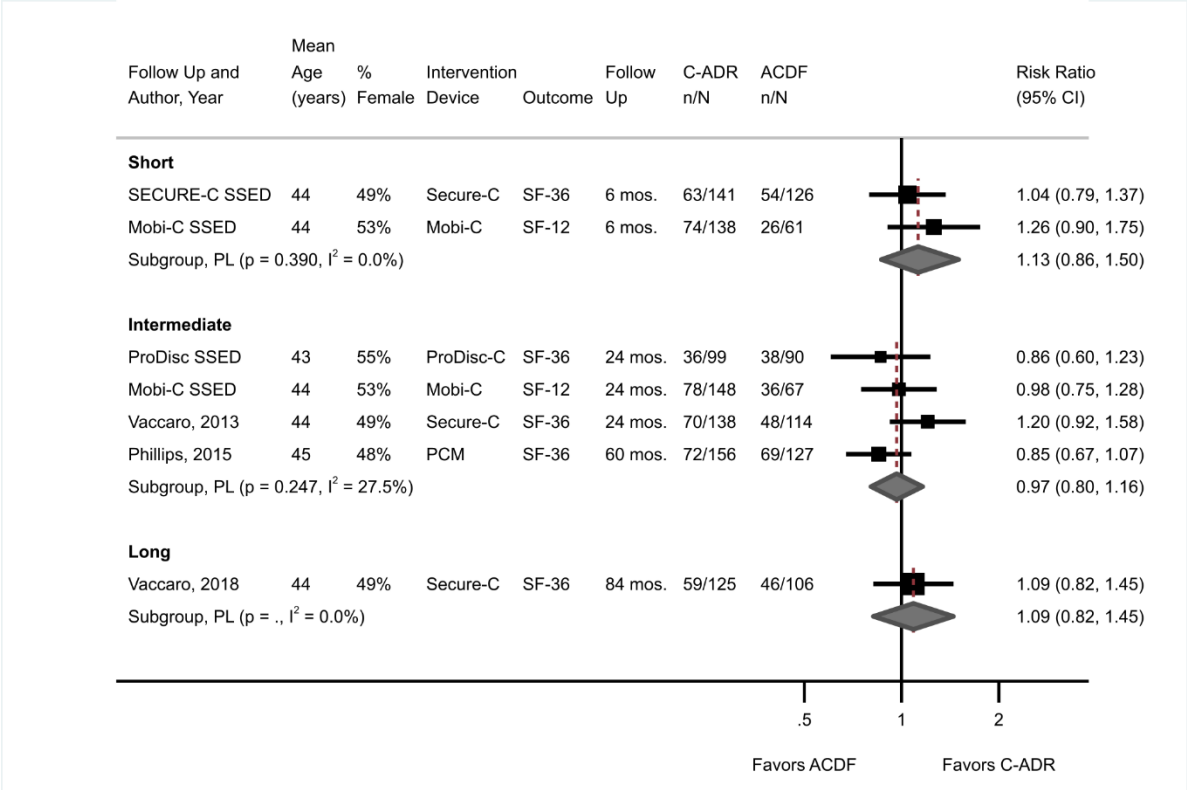
Figure 14. SF-36 or SF-12 PCS success (≥15-point improvement): Comparison of C-ADR with ACDF (1-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PCS = Physical Component Score; PL = profile likelihood; SF-12 = Short Form-12 questionnaire; SF-36 = Short Form-36 questionnaire; SSED = Summary of Safety and Effectiveness Data (FDA).

3. Results

Figure 15. SF-36 or SF-12 MCS success (≥15-point improvement): Comparison of C-ADR with ACDF (1-level interventions)

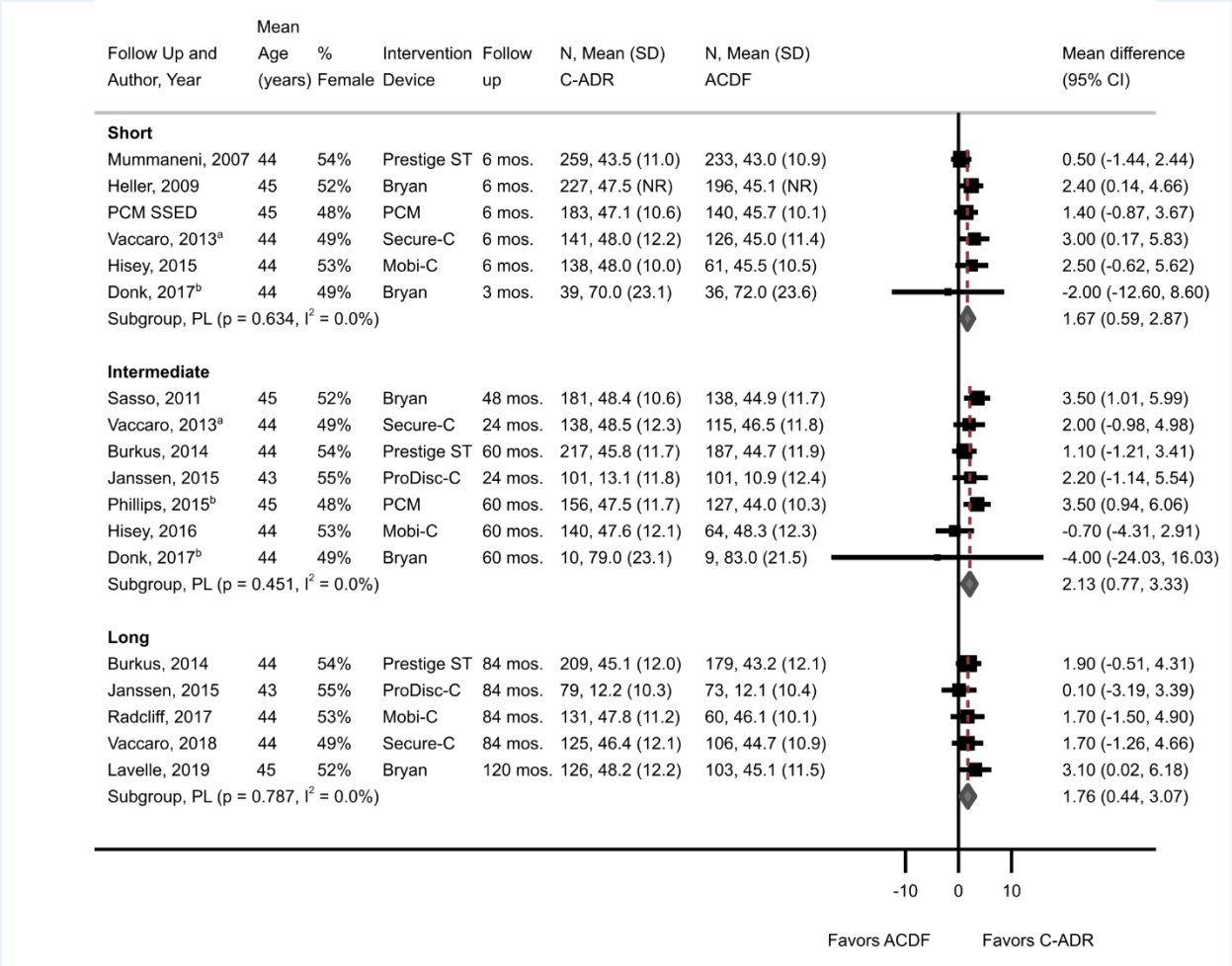


ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; MCS = Mental Component Score; mos. = months; PL = profile likelihood; SF-12 = Short Form-12 questionnaire; SF-36 = Short Form-36 questionnaire; SSED = Summary of Safety and Effectiveness Data (FDA).

Seven RCTs (N=2,368) (in 14 publications)[58,67,73,75,76,79,82,84,90,93-96,112] that compared C-ADR with ACDF reported SF-36/12 PCS and MCS scores (0-100 scale). There were no differences between C-ADR and ACDF in PCS scores (**Figure 16**) as estimates were below the threshold for a small effect in the short-term (6 RCTs, N=1,779, MD 1.67, 95% CI 0.59 to 2.87, I^2=0%), intermediate term (7 RCTs, N=1,684, MD 2.13, 95% CI 0.77 to 3.33, I^2=0%) or long-term (5 RCTs, N=1,191, MD 1.76, 95% CI 0.44 to 3.07, I^2=0%). Similarly, there were no differences between C-ADR and ACDF in MCS scores (**Figure 17**) as estimates were below the threshold for a small effect in the short-term (6 RCTs, N=1,779, MD 1.14, 95% CI -0.14 to 2.17, I^2=0%), intermediate term (7 RCTs, N=1,814, MD 0.83, 95% CI -0.75 to 2.41, I^2=32.2%) and long-term (3 RCTs, N=574, MD 0.64, 95% CI -1.47 to 2.82, I^2=0%). Effect estimates for PCS and MCS did not differ following the exclusion of one trial with unclear samples sizes.[67] No studies were rated high risk of bias.

3. Results

Figure 16. SF-36 or SF-12 PCS scores (0-100): Comparison of C-ADR with ACDF (1-level interventions)

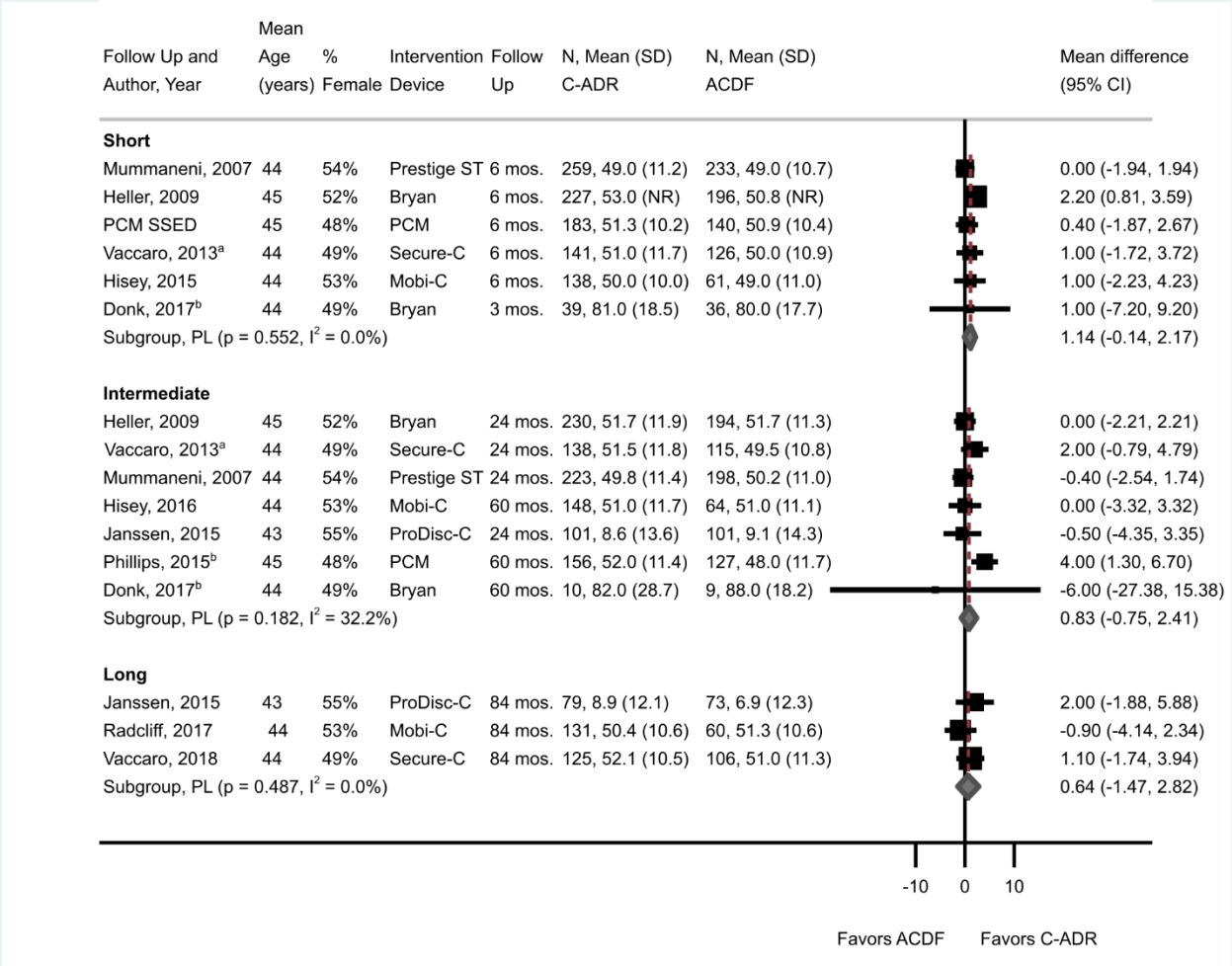


ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PCS = Physical Component Score; PL = profile likelihood; SF-12 = Short Form-12 questionnaire; SF-36 = Short Form-36 questionnaire; SD = standard deviation; SSED = Summary of Safety and Effectiveness Data (FDA).

[a] n/N obtained from the SECURE-C SSED.

[b] Scores estimated from graphs in article.

Figure 17. SF-36 or SF-12 MCS scores (0-100): Comparison of C-ADR with ACDF (1-level interventions)



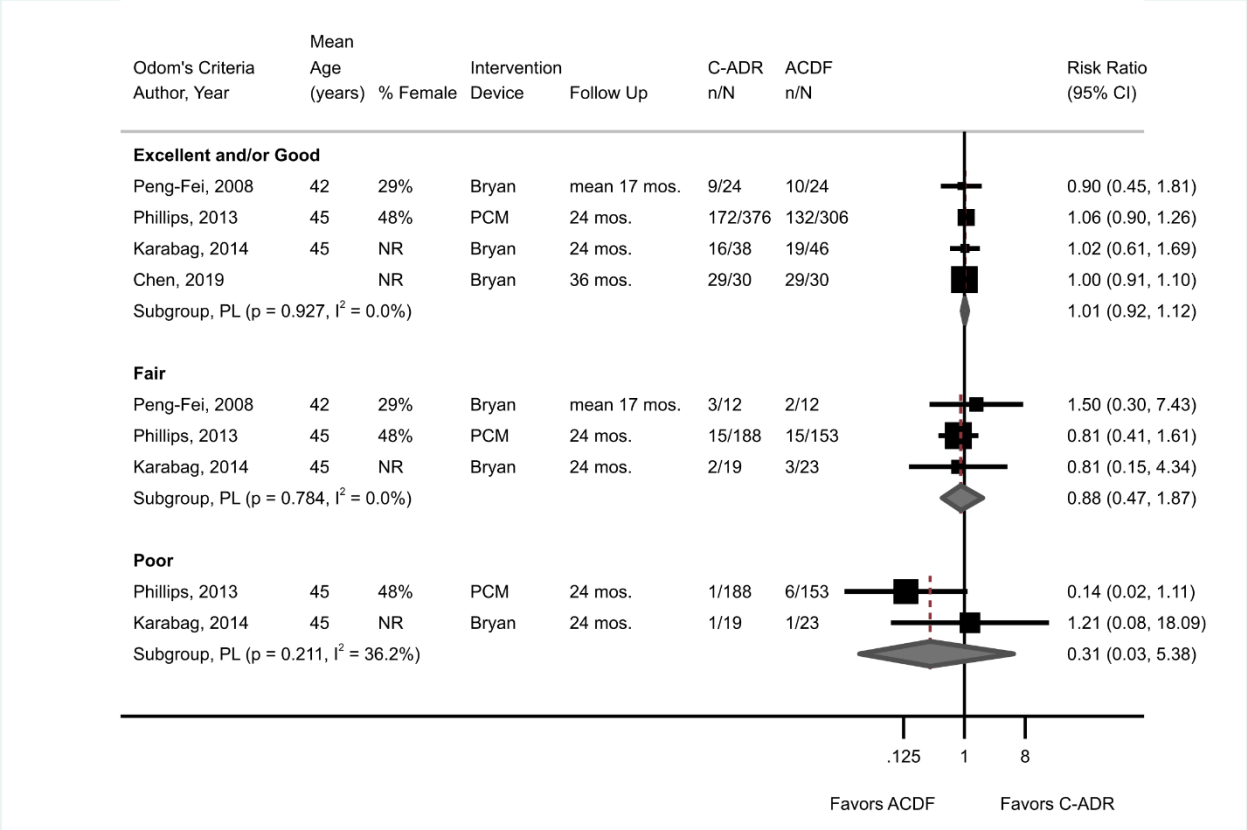
ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; MCS = Mental Component Score; mos. = months; PL = profile likelihood; SF-12 = Short Form-12 questionnaire; SF-36 = Short Form-36 questionnaire; SD = standard deviation; SSED = Summary of Safety and Effectiveness Data (FDA).
[a] n/N obtained from the SECURE-C SSED.
[b] Scores estimated from graphs in article.

3.9.3.1.3.2.3 Odom's Criteria

Four RCTs (N=553)[59,80,89,91] used Odom's Criteria to categorize overall improvement as excellent (i.e., all pre-operative symptoms relieved, abnormal findings improved), good (i.e., minimal persistence of symptoms, abnormal findings unchanged or improved), fair (i.e., definite relief of some symptoms, others unchanged or slightly improved) or poor (i.e., symptoms and signs unchanged or exacerbated). There were no differences between single-level C-ADR and ACDF in the likelihood of having excellent or good results based on Odom's criteria (4 RCTs, N=847, 48.3% vs. 46.8%, RR 1.01, 95% CI 0.92 to 1.12, I^2=0%) at intermediate term.[59,80,89,91] However, three of the RCTs (all small) were rated high risk of bias,[59,80,89] while the one large RCT was rated moderate risk of bias.[91] Based on the highest quality trial, there was no difference between procedures in the likelihood of having excellent or good improvement (1 RCT, N=682, 45.7% vs. 43.1%)[91] (**Figure 18**). In one prospective NRSI IDE study using propensity-matched historical controls, there was no difference between C-ADR and ACDF in the likelihood of

having excellent or good results using Odom's criteria at 24 months (N=301, 90.5% vs. 79.9%).[116]

Figure 18. Odom's Criteria: Comparison of C-ADR with ACDF (1-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PL = profile likelihood.

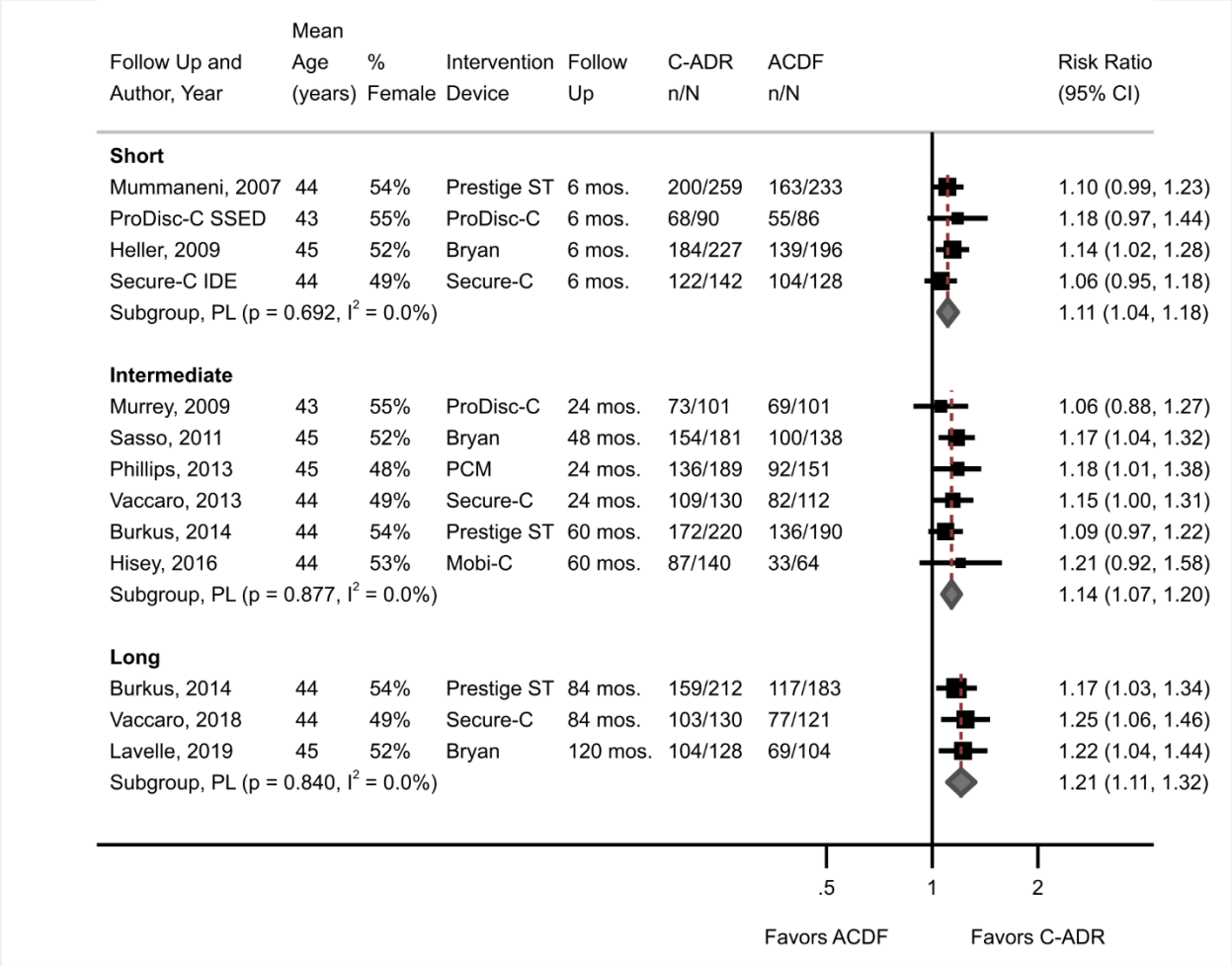
3.9.3.1.3.3 Overall Success (Composite)

The FDA IDE trials were required to report overall success, a composite outcome for six RCTs (N=2,271) (in 11 publications)[58,73,76,82,84,85,91,94,96,114,115] that included a threshold of ≥15-point NDI improvement (0-50 scale) from baseline, improvement or maintenance of neurologic status, no serious adverse events and no additional surgical procedures that might be considered "failure" (e.g., removal, revision, supplemental fixation). In participants with single-level interventions, effect estimates were below the threshold for a small effect and classified as no difference in overall success comparing C-ADR with ACDF in the short term (4 RCTs, N=1,361, 79.9% vs. 71.7%, RR 1.11, 95% CI 1.04 to 1.18, I^2=0%)[73,84,114,115] and intermediate term (6 RCTs, N=1,717, 76.1% vs. 67.7%, RR 1.14, 95% CI 1.07 to 1.20, I^2=0%);[58,76,85,91,94,96] but a slightly increased likelihood of overall success favoring C-ADR was seen long term (3 RCTs, N=878, 76.1% vs. 67.7%, RR 1.21, 95% CI 1.11 to 1.32, I^2=0%)[58,82,96] (**Figure 19**). In one prospective NRSI IDE study using propensity-matched historical controls, there was no difference between C-ADR and ACDF in overall response (same definition as in RCTs) at 24 months (N=301, 86.8% vs. 79.3%, p=0.265).[116]

One of the above trials reported overall success at 84 months using a different criterion for NDI (improvement in NDI score ≥30 points if preoperative score ≥60 or improvement of ≥50% if preoperative score <60) and included an additional requirement for radiographic success, and

was not included in the meta-analysis at long term; there was no difference between C-ADR and ACDF using this criteria (N=166, 55.2% vs. 50.0%, RR 1.10, 95% CI 0.80 to 1.52).[93]

Figure 19. Overall success: Comparison of C-ADR with ACDF (1-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PL = profile likelihood; SSED = Summary of Safety and Effectiveness Data (FDA).

3.9.3.1.3.4 Quality of Life

None of the included studies reported on quality-of-life measures.

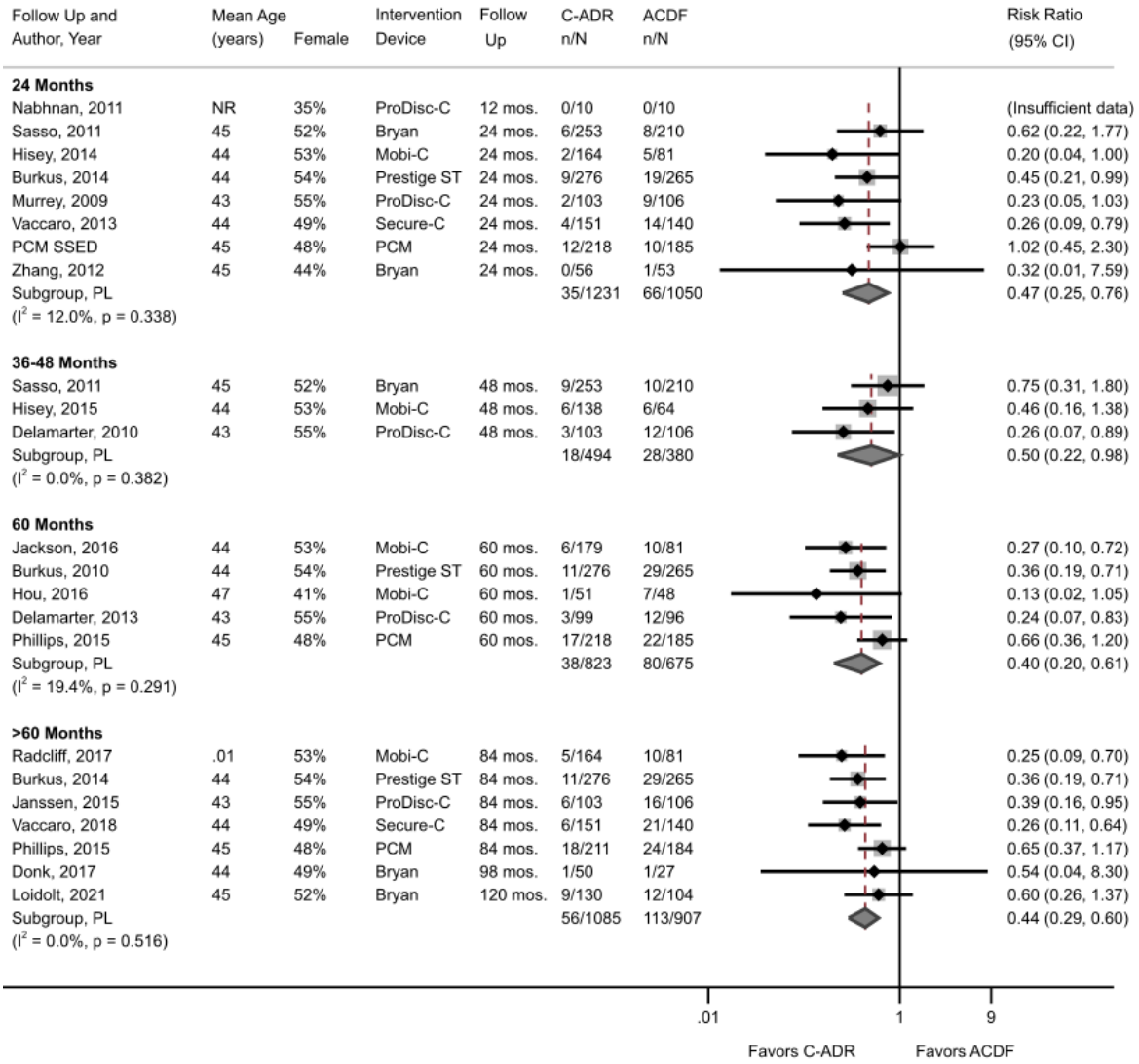
3.9.3.1.3.5 Reoperation and subsequent surgery

There was high-strength evidence that C-ADR was associated with substantially lower likelihood of reoperation at the index level versus ACDF (SOE: High).

Reoperation including any additional procedure at the index level was substantially less frequent with C-ADR versus ACDF for single-level disease at all time points reported in RCTs including short term up to 24 months (8 RCTs, N=2,281, 2.8% vs. 6.3%, RR 0.47, 95% CI 0.25 to 0.76, I^2=12%; RD 3 per 100 participants, 95% CI 1 per 100 to 5 per 100 participants)[58,74,85,88,94,96,98,112] and long term from 84 to 120 months (7 RCTs, N=1,992, 5.2% vs. 12.5%, RR 0.44, 95% CI 0.29 to 0.60, I^2=0%; RD 7 per 100 participants, 95% CI 4 per 100 to 9 per 100 participants)[58,67,79,83,90,93,95] (**Figure 20**).

3. Results

Figure 20. Reoperation at the index level: Comparison of C-ADR with ACDF (1-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PL = profile likelihood; SSED = Summary of Safety and Effectiveness Data (FDA).

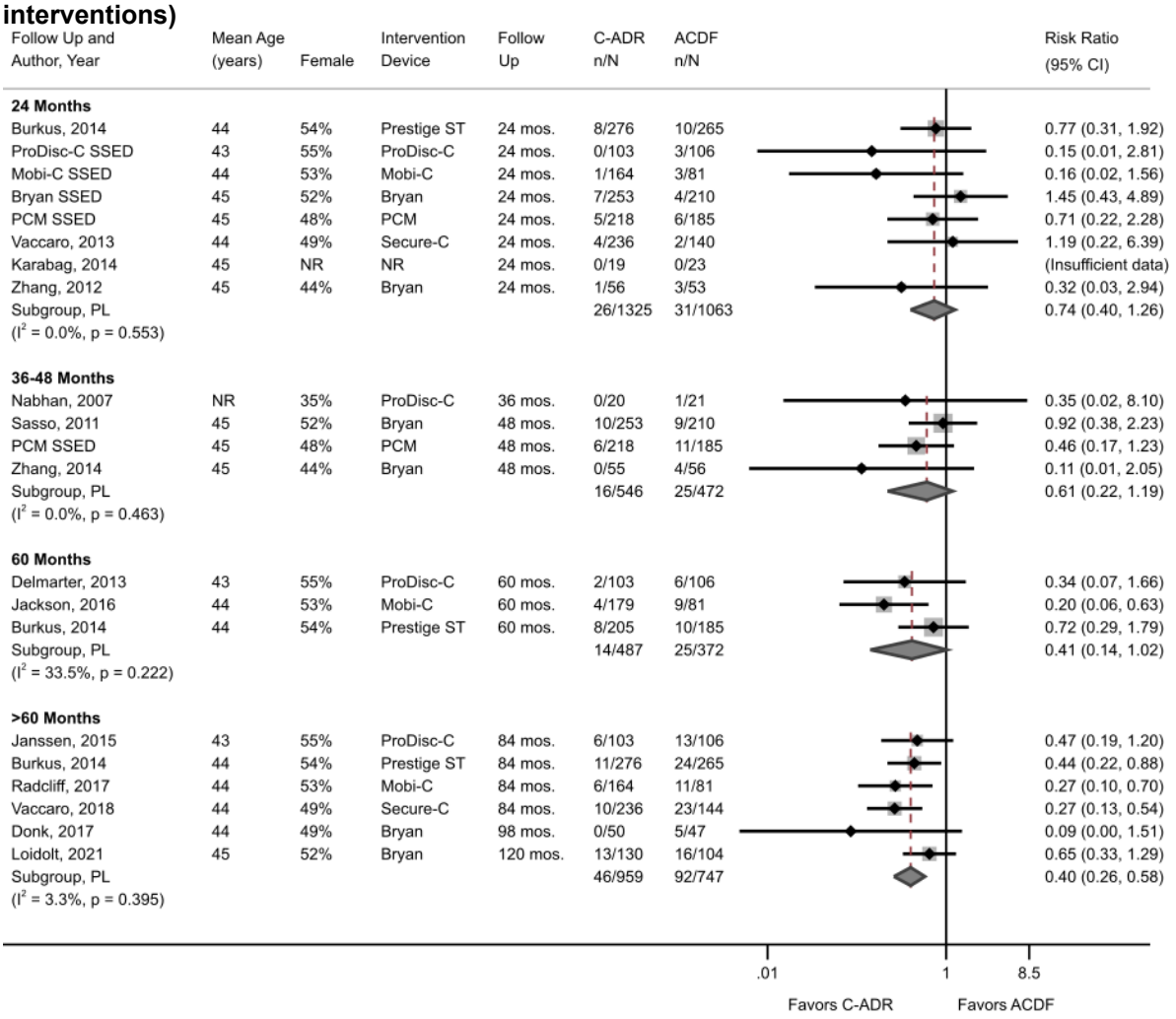
One prospective NRSI IDE study of C-ADR using historical ACDF controls found no difference in index-level reoperation up to 24 months (N=349, 1.9% vs. 4.8%, RR 0.39, 95% CI 0.11 to 1.43).[116]

Reoperation across two NRSIs was less common than that reported in RCTs. No difference in 30-day reoperation was seen in one NRSI (1.2% vs. 0.4%, adjusted OR 0.60, 95% CI 0.14 to 2.56).[107] Another NRSI reported that reoperation was less common following C-ADR within 90 days of index surgery compared with ACDF (2.04% vs. 3.35%, adjusted OR 0.63, 95% CI 0.44 to 0.92) but no difference between C-ADR and ACDF longer-term up to 5 years (adjusted HR 0.86, 95% CI 0.60 to 1.23).[106] While overall reoperation rates were lower in these database NRSIs, it is possible the RCTs, particularly IDE trials may provide more accurate detail regarding specific indications.

Subsequent surgery rates at adjacent levels were similar between C-ADR and ACDF at up to 24 months[58,80,96,98,110-112,114] and between 36 and 48 months (including after exclusion of one trial

rated high risk of bias[99])[87,94,99,112] but was substantially less likely with C-ADR versus ACDF at 60 months (3 RCTs, N=859, 2.9% vs. 6.8%, RR 0.41, 95% CI 0.14 to 1.02, I^2=33.5%)[58,66,78] and at the longest follow-ups from 84 to 120 months (6 RCTs, N=1706, 4.8% vs. 12.3%, RR 0.40, 95% CI 0.26 to 5.8, I^2=3.3%).[58,67,79,83,93,95] However, estimates were somewhat imprecise (**Figure 21**). Also, across trials, indications for operation at adjacent levels were not consistently described.

Figure 21. Subsequent surgery at adjacent levels: Comparison of C-ADR versus ACDF (1-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PL = profile likelihood; SSED = Summary of Safety and Effectiveness Data (FDA).

3.9.3.1.3.6 Harms

All 15 RCTs that evaluated C-ADR and ACDF for single-level disease provided information on adverse events and harms up to 120 months followup.[58,59,67,74,77,80,85,87,88,90,94,96,98,99,112] Information on harms from four NRSIs was used to complement that from RCTs.[104,106,107,116]

3.9.3.1.3.6.1 Neurologic deficit

There was low-strength evidence of no differences in the likelihood of neurological events or deficits between C-ADR and ACDF at short, intermediate, or long-term (SOE: Low).

Reporting of neurological events varied across RCT publications. Three trials assessed use of the Bryan IDE trial at different times;[56,83,94] one IDE trial evaluated Mobi-C.[93] One trial[56] described specific, observed neurological events as acute neurological changes, while other trials used various general terms to describe neurologic events (e.g., new deficit, neurological failure, neurological AE). The timing of events following surgery was also not clearly reported. Thus, reported proportions of participants who experienced neurological events varied substantially across RCTs, however there were no differences between C-ADR and ACDF at 0 to 24 months (3.3% vs. 3.2%),[56] between 24 and 48 months (0% vs. 1.0%, WHO grade 3 or 4),[94] up to 84 months (11.4 % vs. 11.5%)[93] or up to 120 months (any: 43.1% vs. 43.8%; WHO grade 3 or 4: 4.5% vs. 6.9%).[83] One prospective NRSI IDE study of C-ADR that used propensity-matched historical ACDF controls reported no differences in serious device- or procedure-related neurological adverse events between C-ADR and ACDF (1.3% vs. 1.6%) through 24 months.[116] The same trial study also reported fewer C-ADR participants experienced neurological decrease from baseline versus ACDF (6.7% vs. 12.8%, RR 0.52, 95% CI 0.25 to 1.07) but results were imprecise.

3.9.3.1.3.6.2 Death

There was inadequate evidence to draw conclusions on the likelihood of death in participants undergoing C-ADV versus ACDF (SOE: Insufficient).

Death was uncommon (<3%) in RCTs and NRSIs, with no reported differences between C-ADR and ACDF. Across RCTs, no deaths were directly attributed to either procedure, however cause of death was not reported in many trials. For C-ADR from 0 to 24 months, three of the four deaths were attributed to myocardial infarction or cardiac arrest in one trial;[58] the cause of the fourth death was not reported in another trial.[96] No deaths were observed in one trial.[74] At followup from 0 to 36 months, one C-ADR participant died of a severe subarachnoid hemorrhage at 6 weeks (relationship to procedures was not stated)[87] and one death in the ACDF group attributed to a motor vehicle accident was observed in another trial.[56] There was no difference in mortality between procedures at 84 months (1 RCT, N=541, 0.9% vs. 2.2%, RR 0.38, 95% CI 0.08 to 1.96)[58] or at 120 months (1 RCT, N=232, 1.4% vs. 2.4%, RR 0.54, 95% CI 0.09 to 3.18),[83] however estimates were imprecise. Findings from one large administrative data NRSI[106] reinforce that death was rare for C-ADR (0%) and ACDF (0.18%) and that there was no difference between procedures in the likelihood of mortality. One death occurred in the C-ADR group in one NRSI IDE study using historical controls up to 24 months[116] (**Appendix C**).

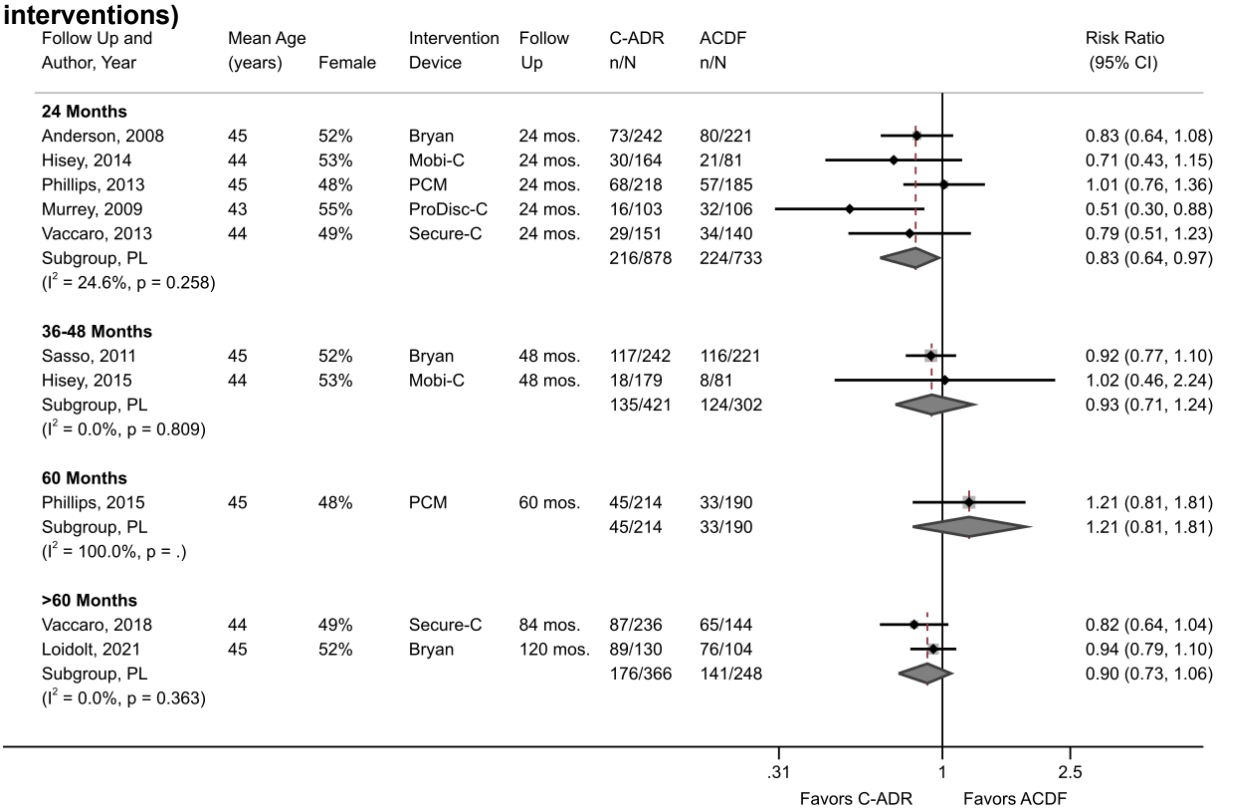
3.9.3.1.3.6.3 Serious Adverse Events

There was low-strength evidence that C-ADR was associated with a slightly lower likelihood of any serious adverse event in the short term versus ACDF (SOE: Low); there was also low-strength of no differences in the likelihood of experiencing a serious adverse event at greater than 24 months (SOE: Low).

Serious adverse event definitions and types of events varied across RCTs, but often included events that were life threatening, required medical intervention, or resulted in a permanent disability or death. Timing of events was not reported. Events related to participant factors such as comorbidities (e.g., underlying cardiovascular disease) would likely not be different between procedures. C-ADR was associated with a slightly lower likelihood of experiencing a serious adverse event up to 24 months across IDE trials (5 RCTs, N=1,611, 24.6% vs. 30.6%, RR 0.83, 95% CI 0.64 to 0.97, I^2= 24.6%)[56,74,85,91,96] compared with ACDF, however across fewer trials at other times, no differences between procedures was seen (**Figure 22**). No difference in the

likelihood of experiencing a serious AEs was seen between C-ADR and ACDF (N=349, 9.4% vs. 14.8%, RR 1.97, 95% CI 0.88 to 4.37) in one NRSI IDE study using historical controls up to 24 months.

Figure 22. Any serious adverse events (author defined): Comparison of C-ADR with ACDF (1-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PL = profile likelihood.

Dysphagia was reported by six RCTs (N=1,965) (in 8 publications),[56,58,66,67,74,79,83,96] but the severity was unclear in most cases. One trial (N=463) reported no cases of WHO grade 3 or 4 dysphagia in any participant through 24 months followup.[56]

NRSIs based on administrative data suggest that serious adverse events are rare and not different between C-ADR and ACDF. Thrombolic event rates (DVT and/or PE) were similar between C-ADR (range 0.07% to 0.19%) and ACDF (0.10% to 0.11%) as reported by two large NRSIs.[104,106] One NRSI[106] reported rates of vertebral artery injury and dural tear of less than 1% in for each procedure. (Appendix). One NRSI reported low risk of dysphagia (0% vs. 0.13%)[107] but did not report dysphagia severity. Dysphagia was more common in C-ADR participants versus ACDF participants (9.4% vs. 6.3%) but severity was not described in one prospective NRSI IDE study using historical ACDF controls.[116]

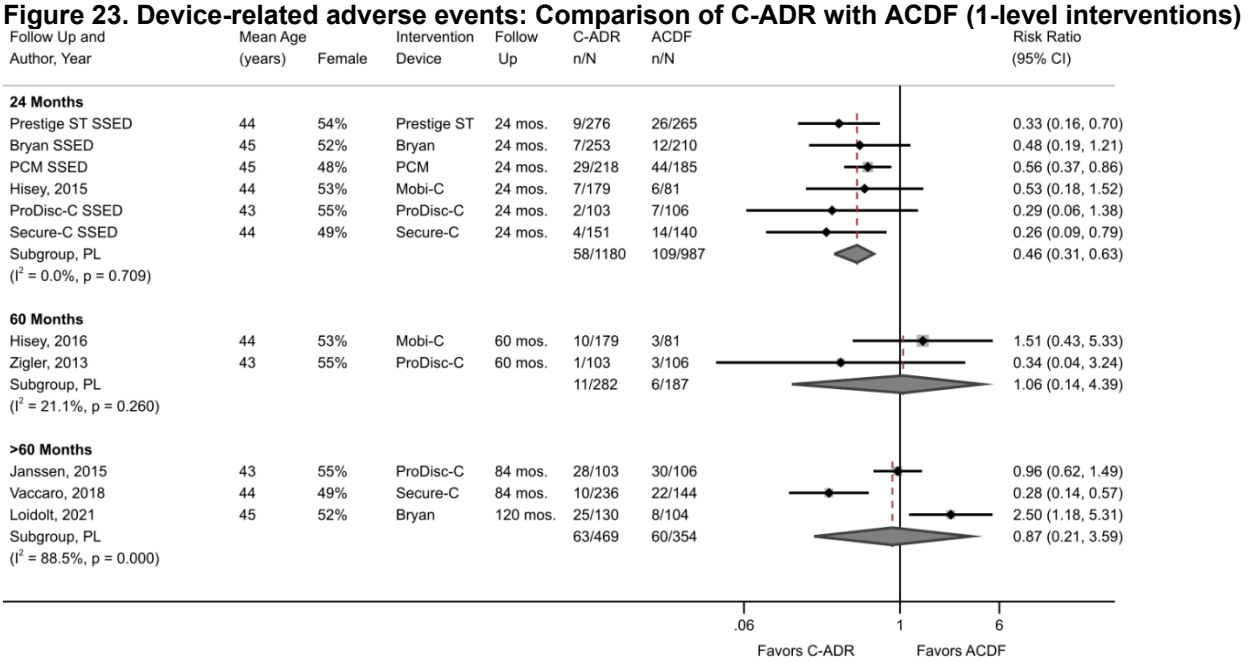
3.9.3.1.3.6.4 Heterotopic Ossification

Grade 3 or 4 heterotopic ossification (HO) may be of concern with C-ADR. Across 4 RCTs, (N=398 for C-ADR arm), 6.3 percent of participants developed Grade 3 or 4 HO.[59,76,90,98]

3.9.3.1.3.6.5 Device-related Adverse Events

Device-related adverse event definitions, types of events and adjudication varied across RCTs. Some trials included a range of events such as adjacent -level degenerative joint changes, headache as well as neurological events. Some device-related events may only occur with C-ADR, others may only occur with ACDF (e.g., nonunion). Some events may not be persistent or serious (e.g., superficial wound infection, dysphagia). C-ADR was associated with substantially lower likelihood of device-related events at 24 months (6 RCTs, N=2,167, 4.9% vs. 11%, RR 0.46, 95% CI 0.31 to 0.63, I^2=0%).[75,111-115] No difference was seen across two trials at 60 months,[76,100] but results across three trials at >60 months[79,83,95] were inconsistent (**Figure 23**).

Figure 23. Device-related adverse events: Comparison of C-ADR with ACDF (1-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PL = profile likelihood; SSED = Summary of Safety and Effectiveness Data (FDA).

3.9.3.1.3.6.6 Differential effectiveness (HTE)

None of the included trials that compared single-level C-ADR and ACDF interventions reported differential effectiveness based on patient or other characteristics.

3.9.3.2 Two-level C-ADR versus ACDF

Four RCTs (N=872) (in 11 publications)[61,63,64,69-71,78,81,92,93,97] compared two-level C-ADR and ACDF, including two FDA IDE trials (in 9 publications)[63,64,69-71,78,81,92,93] and two non-IDE trials.[61,97] One FDA IDE NRSI[102] compared a novel polyetheretherketone (PEEK)-on-ceramic C-ADR with propensity score-matched historical ACDF controls (structural allograft and plate) from a multicenter RCT initiated in the mid-2000s that was not referenced.

3.9.3.2.1 Fusion

Two RCTs (N=727) (across 4 publications) that compared two-level C-ADR and ACDF procedures reported fusion success in their ACDF arms.[69,81,92,93] No trials reported short-term fusion success. Two RCTs (N=243) reported intermediate-term fusion success in 92.5 percent (Range: 90.5% to 94.0%) of participants.[81,92] Two RCTs (N=196) reported long-term fusion

success in 92.6 percent (Range: 90.9% to 93.8%) of participants.[69,93] One IDE NRSI[102] comparing a novel C-ADR versus historical ACDF controls reported pseudarthrosis in 6.5 percent of the ACDF group.

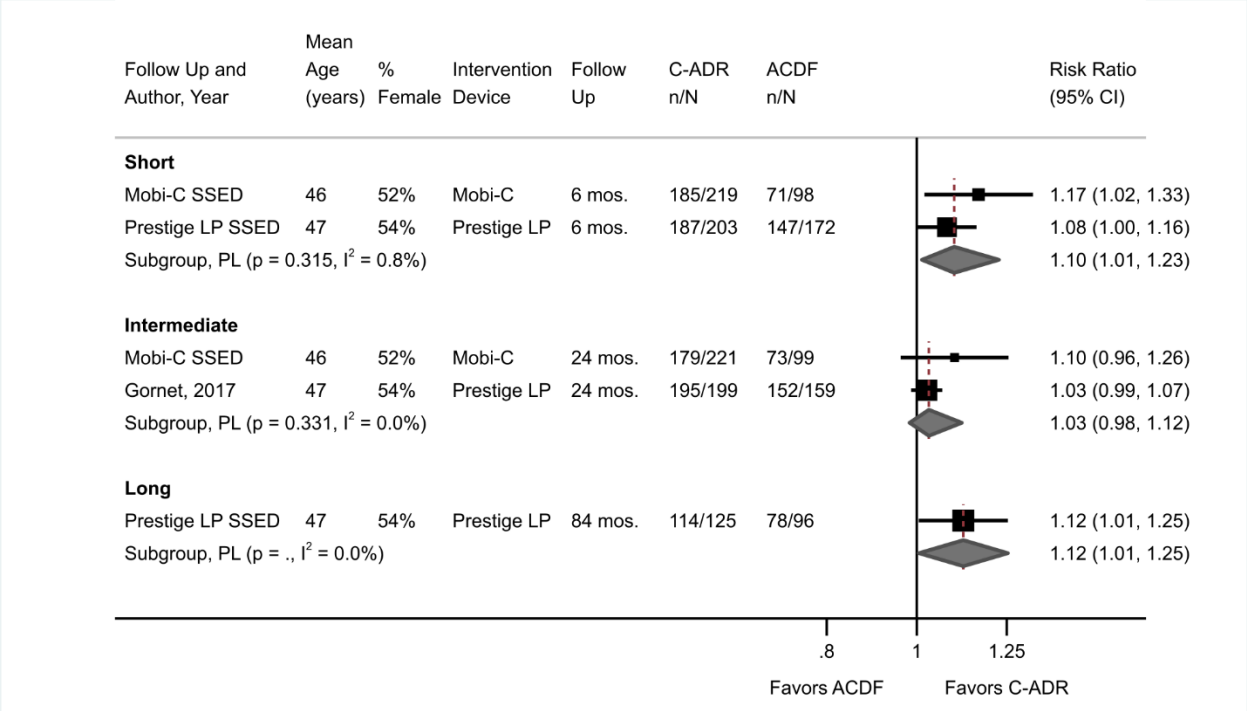
3.9.3.2.2 Pain

3.9.3.2.2.1 Neck pain

There was moderate-strength evidence of no difference between C-ADR and ACDF on neck pain (SOE: Moderate).

Two RCTs (in 3 publications) (N=727)[70,117,118] that compared C-ADR with ACDF reported ***neck pain success*** (response) defined as postoperative ≥20-point improvement on VAS (0-100 scale). In participants having two-level interventions there were no differences in likelihood of neck pain success between C-ADR and ACDF in the short term (2 RCTs, N=692, 88% vs. 80.7%, RR 1.10, 95% CI 1.01 to 1.23, I^2= 0.8%),[117,118] intermediate term (2 RCTs, N = 678, 89.0% vs. 87.2%, RR 1.03, 95% CI 0.98 to 1.12, I^2=0%)[70,117] and long term (1 RCT, N=221, 91.2% vs. 81.3%, RR 1.12, 95% CI 1.01 to 1.25)[118] (**Figure 24**). There was also no difference long term between C-ADR and ACDF in the trial using a threshold of ≥10-point improvement for neck pain success that was not included in the meta-analysis (1 RCT, N=269, 86% vs 77.7%, RR 1.11, 95% CI 0.97 to 1.32).[93]

Figure 24. Neck pain success (≥20-point improvement on VAS): Comparison of C-ADR with ACDF (2-level interventions)

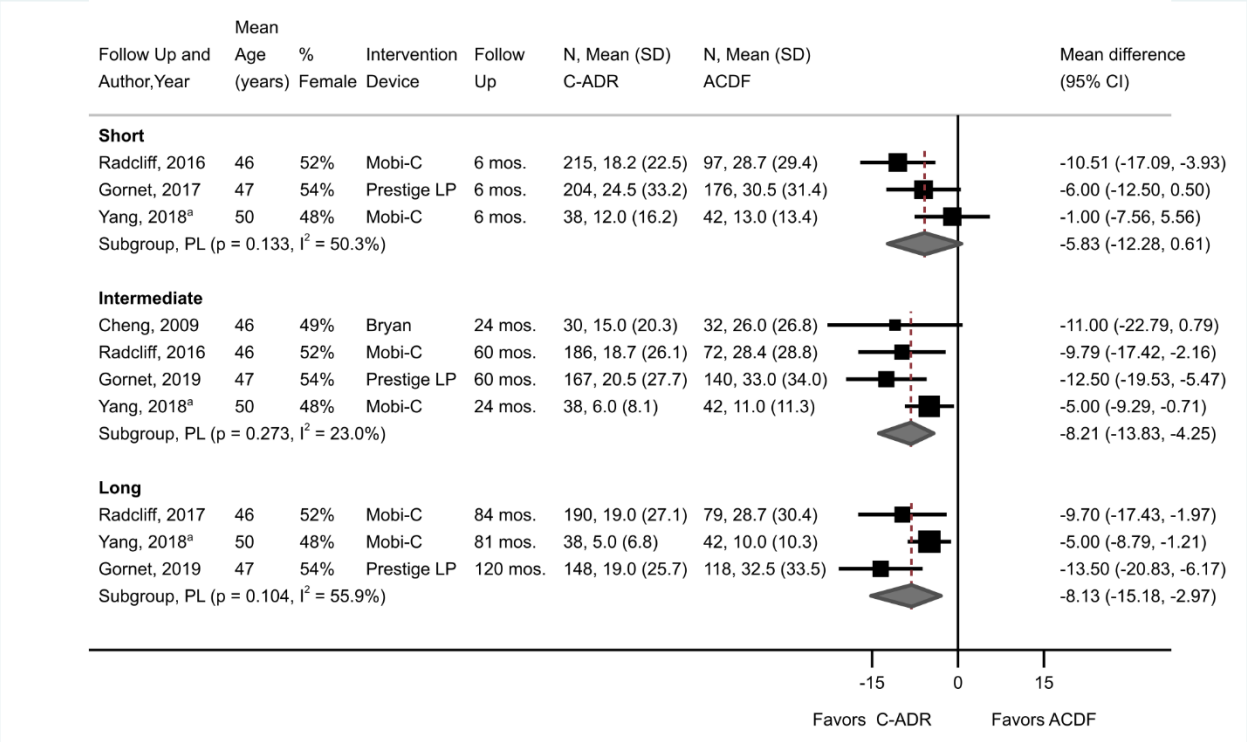


ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PL = profile likelihood; SSED = Summary of Safety and Effectiveness Data (FDA); VAS = visual analog scale.

There was no difference in VAS neck pain scores (0-100 scale) between C-ADR and ACDF short term (3 RCTs, N=764, MD -5.83, 95% CI -12.28 to 0.61, I^2=50.3%).[70,92,97] C-ADR was associated with a small pain improvement versus ACDF in the intermediate term (4 RCTs,

N=707, MD -8.21, 95% CI -13.83 to -4.25, I^2=23%)[61,69,92,97] and long-term (3 RCTs N=615, MD -8.13, 95% CI -15.18 to -2.97, I^2=55.9%)[69,93,97] (**Figure 25**). One IDE NRSI that compared a novel C-ADR versus historical ACDF controls reported no differences in mean VAS neck pain intensity at short- or intermediate term (N=352, 1.8 vs. 2.5 at both times, p>0.10).[102]

Figure 25. Neck pain scores (0-100): Comparison of C-ADR with ACDF (2-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PL = profile likelihood; SD = standard deviation.

[a] Scores estimated from graphs in article.

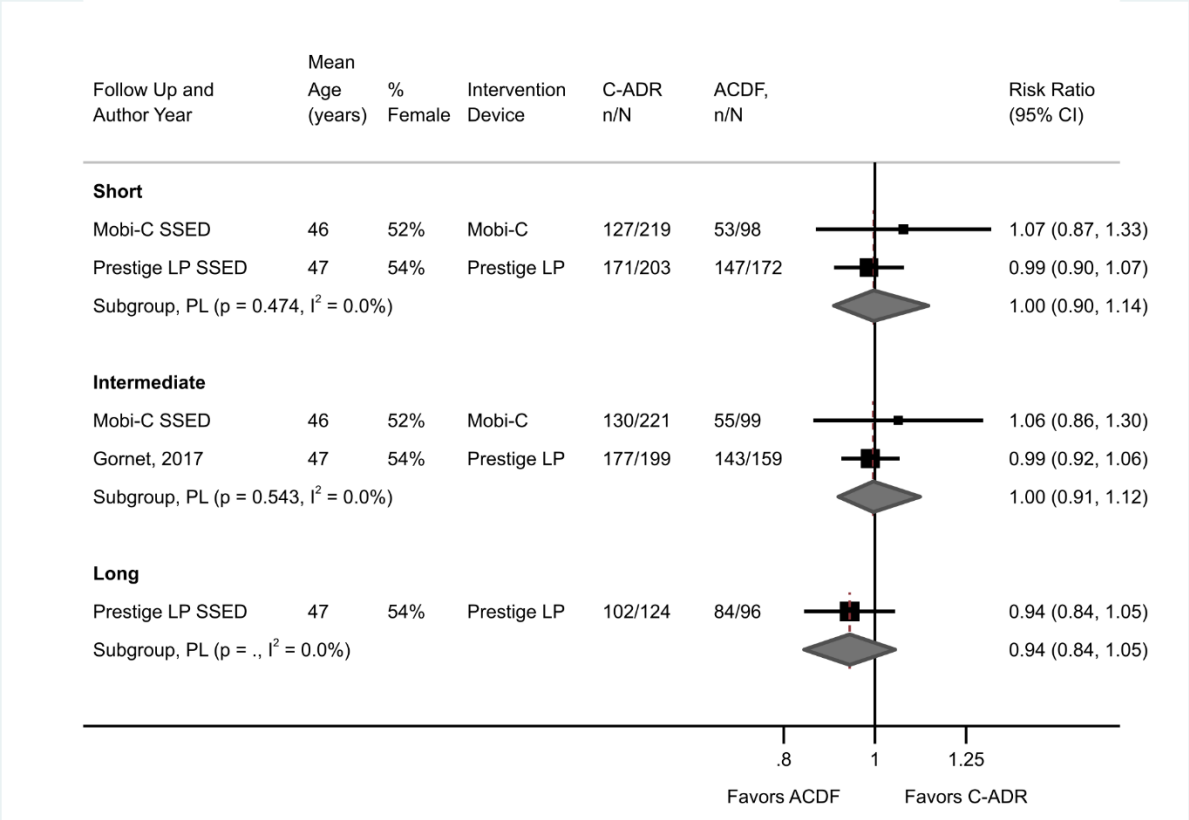
3.9.3.2.2.2 Arm pain

There was moderate-strength evidence of no difference between C-ADR and ACDF on arm pain (SOE: Moderate).

Two RCTs (in 3 publications) (N=727)[70,117,118] that compared C-ADR with ACDF reported arm pain success (response) defined as postoperative ≥20-point improvement on VAS (0-100 scale). Some studies reported arm pain success in both arms. Using conservative estimates (the lower risk ratio), found no difference in likelihood of arm pain success between C-ADR and ACDF at short term (2 RCTs, N=692, 70.6% vs. 74.1%, RR 1.0, 95% CI 0.90 to 1.14, I^2= 0%),[117,118] intermediate term (2 RCTs, N=678, 73.1% vs.76.7%, RR 1.00, 95% CI 0.91 to 1.12, I^2= 0%),[70,117] or long term (1 RCT, N=220, RR 0.94, 95% CI 0.84 to 1.05)[118] (**Figure 26**). Estimates and conclusions using the higher risk ratios from the other arm were similar.

3. Results

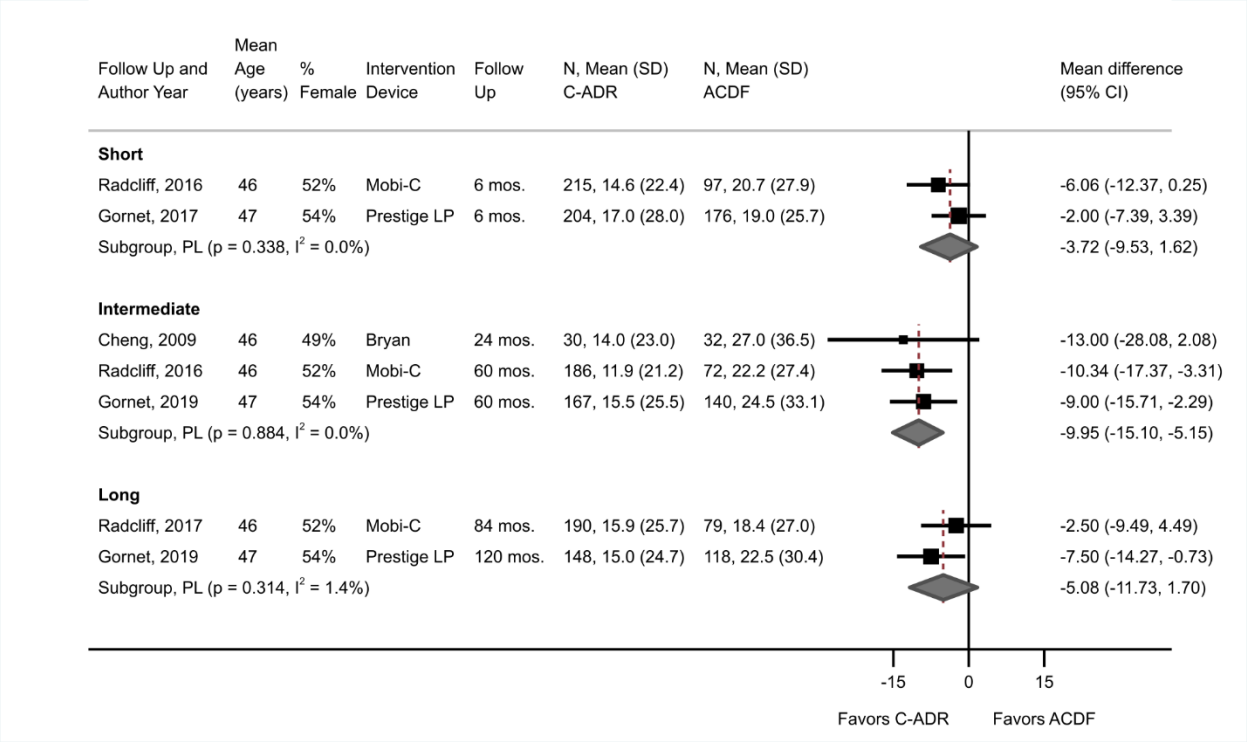
Figure 26. Arm pain success (≥20-point improvement on VAS): Comparison of C-ADR with ACDF (2-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PL = profile likelihood; SSED = Summary of Safety and Effectiveness Data (FDA); VAS = visual analog scale.

Three RCTs (N=792) (in 5 publications)[61,69,70,92,93] reported arm pain scores (0-100). Some trials reported arm pain scores in both arms. Conservative estimates (using the smaller mean differences) are reported here. There was no difference in VAS arm pain scores (0-100 scale) between C-ADR and ACDF in the short term (2 RCTs, N=692, MD -3.72, 95% CI -9.53 to 1.62, I^2=0%).[70,92] C-ADR was associated with a small pain improvement versus ACDF at intermediate term (3 RCTs, N=627, MD -9.95, 95% CI -15.10 to -5.15, I^2=0%)[61,69,92] but not long-term (2 RCTs N=535, MD -5.08, 95% CI -11.73 to 1.70, I^2=1.4%)[69,93] (**Figure 27**). One IDE NRSI (N=352) that compared a novel C-ADR versus ACDF using historical controls reported no differences in mean VAS arm pain intensity at short (1.6 vs. 1.7) or intermediate term (1.8 vs. 1.6).[102]

Figure 27. Arm pain scores (0-100): Comparison of C-ADR with ACDF (2-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PL = profile likelihood; SD = standard deviation; SSED = Summary of Safety and Effectiveness Data (FDA).

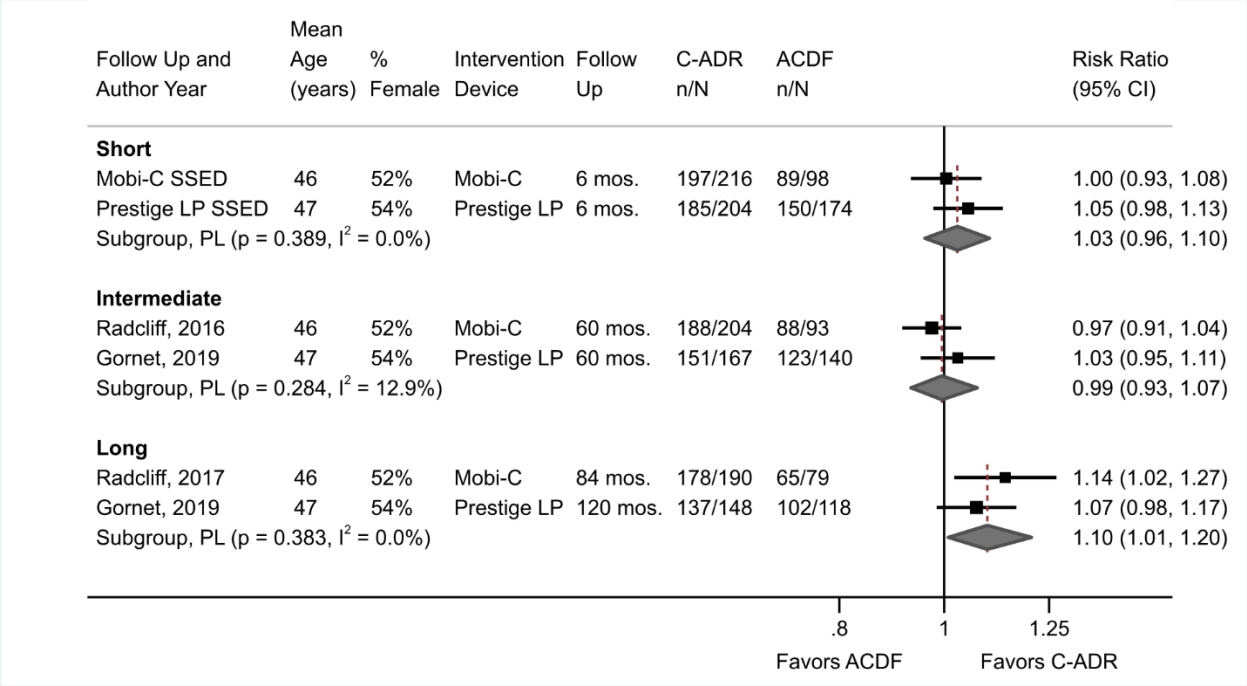
3.9.3.2.3 Function

3.9.3.2.3.1 Neurologic Function

There was moderate-strength evidence of no difference between C-ADR and ACDF on neurologic function (SOE: Moderate).

Two IDE RCTs (N=727) (in 5 publications)[69,92,93,117,118] that compared C-ADR with ACDF reported neurologic success (response), defined as maintenance or improvement (compared with preoperative status) in motor function, sensory function, and deep tendon reflexes. In participants with two-level interventions, there was no difference in likelihood of neurologic success between C-ADR and ACDF at short-term (2 RCTs, N=692, 91.0% vs. 87.9%, RR 1.03, 95% CI 0.96 to 1.10, I^2= 0%),[117,118] intermediate term (2 RCTs, N=604, 91.4% vs. 90.6%, I^2=12.9%)[69,92] or long term (2 RCTs, N=535, 93.2% vs. 84.8%, I^2= 0%)[69,93] (**Figure 28**). The likelihood of neurological success, based on motor, sensory, and myelopathic gait assessments, was similar for C-ADR and ACDF in one IDE NRSI (N=352, 100% vs. 97.7%).[102]

Figure 28. Neurologic success: Comparison of C-ADR with ACDF (2-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PL = profile likelihood; SSED = Summary of Safety and Effectiveness Data (FDA).

Mean JOA scores (0-17 scale) were similar following C-ADR and ACDF at short term (6 months, 15.2 vs. 14.9, p>0.05), intermediate term (15.4 vs. 15.3, p>0.05), and long term (81 months, 15.4 vs. 15.2, p>0.05) in one RCT (N=96).[97]

3.9.3.2.3.2 General Function

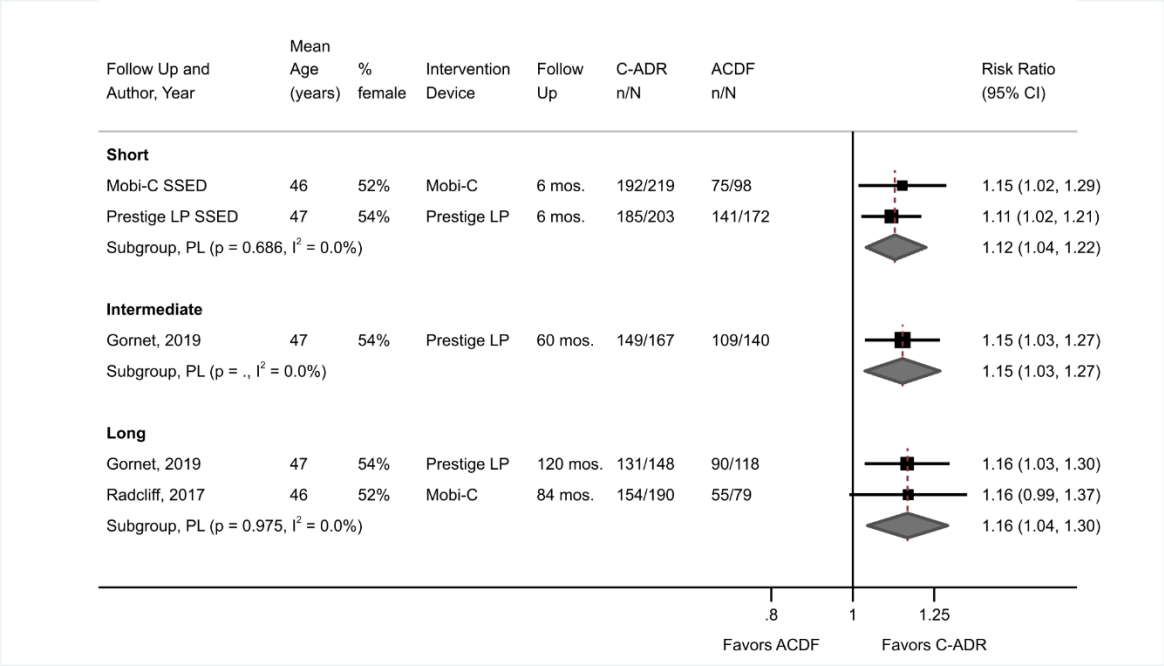
There was moderate-strength evidence of no difference between C-ADR and ACDF on general function (SOE: Moderate).

3.9.3.2.3.2.1 NDI

Two IDE RCTs (N=727) (in 4 publications)[69,93,117,118] and one IDE NRSI (N=352)[102] that compared C-ADR with ACDF reported NDI success defined as postoperative NDI score improvement of ≥15 points from baseline. One trial defined NDI success as improvement of ≥30 points from baseline and was not included in the meta-analysis.[64] Based on the threshold of ≥15 points from baseline, there were no differences between C-ADR and ACDF (although statistically significant, the difference between treatments is too small to be meaningful; see **Table 2**) at short term (2 RCTs, N=692, 89.3% vs. 80.0%, RR 1.12, 95% CI 1.04 to 1.22, I^2= 0%),[117,118] intermediate term (1 RCT, N=307, 89.2 % vs. 77.9%, RR 1.15, 95% CI 1.03 to 1.27)[69] and long term (2 RCTs, N=535, 84.3% vs. 73.6%, RR 1.16, 95% CI 1.04 to 1.30, I^2= 0%)[69,93] (**Figure 29**). There was no difference in the likelihood of NDI success between C-ADR and ACDF in one IDE NRSI (N=352, 92.3% vs. 85.5%, p>0.05).[102]

3. Results

Figure 29. NDI success (≥15-point improvement): Comparison of C-ADR with ACDF (2-level interventions)

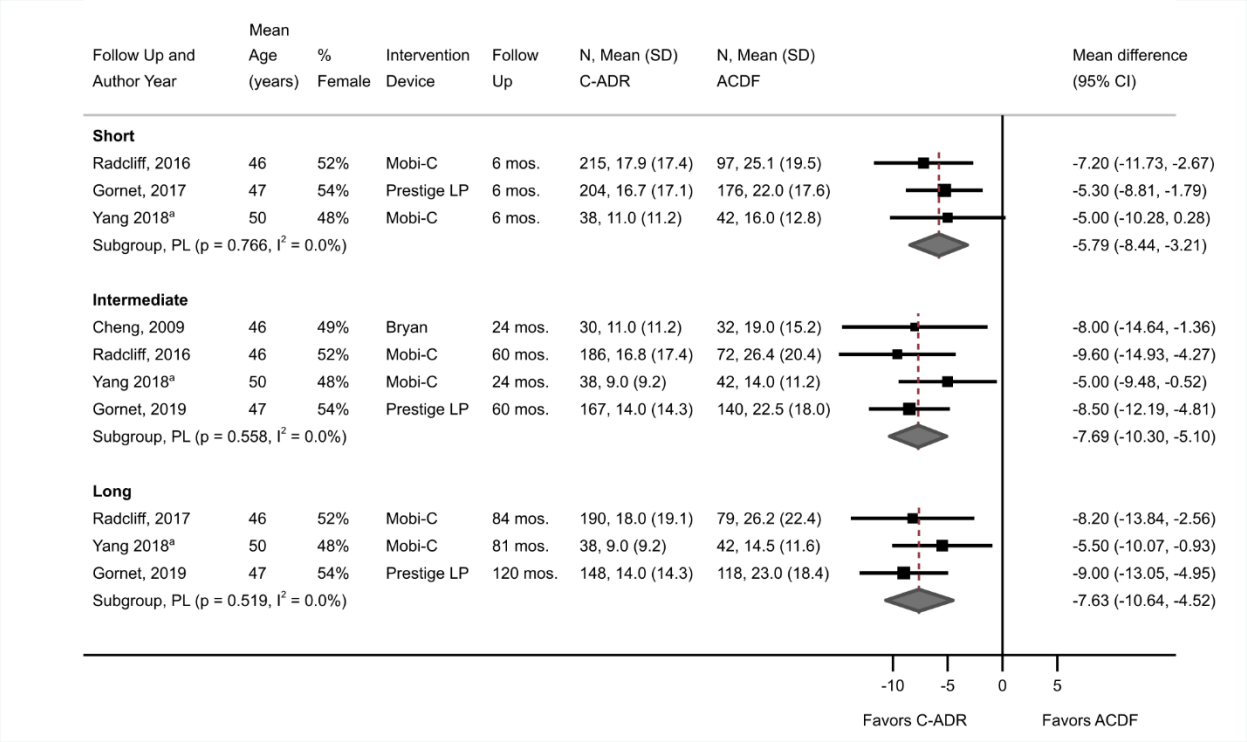


ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; NDI = Neck Disability Index; PL = profile likelihood; SSED = Summary of Safety and Effectiveness Data (FDA).

One RCT that defined NDI success as improvement of ≥30 points from baseline found a moderately higher likelihood of NDI success following C-ADR versus ACDF at intermediate term (1 RCT, N=359, 79.3% vs. 53.4%, RR 1.50, 95% CI 1.21 to 1.86).[64]

Four RCTs (N=872) (in 6 publications)[61,69,70,92,93,97] that compared C-ADR with ACDF reported NDI scores (0-100, higher score, more limitations). C-ADR was associated with a small improvement in function based on NDI scores at short (3 RCTs, N=772, MD -5.79, 95% CI -8.44 to -3.21, I^2=0%),[70,92,97] intermediate (4 RCTs, N=707, MD -7.69, 95% CI -10.30 to -5.10, I^2=0%)[61,69,92,97] and long term (3 RCTS, N=615, MD -7.63, 95% CI -10.64 to -4.52, I^2=0%)[69,93,97] (**Figure 30**).

Figure 30. NDI scores (0-100): Comparison of C-ADR with ACDF (2-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; NDI = Neck Disability Index; PL = profile likelihood; SD = standard deviation.

[a] Scores estimated from graphs in article.

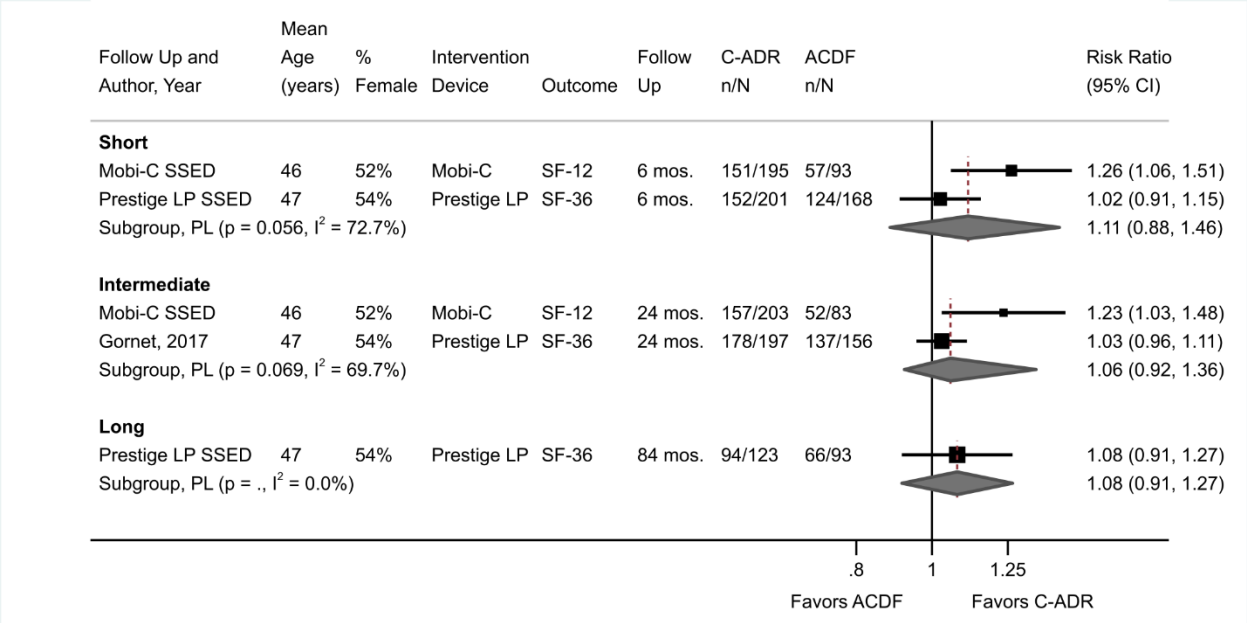
One IDE NRSI (N=352) that compared a novel C-ADR versus historical ACDF controls found that C-ADR was associated with a small improvement in function based on the NDI short term (MD 5.7, means 15.1 vs. 20.8, $p<0.05$); this was not sustained to intermediate term (MD 2.9, means 14.3 vs. 17.2, $p>0.05$).[102]

3.9.3.2.3.2.2 SF-36 PCS and MCS

Two IDE RCTs (N=727) (in 3 publications)[70,117,118] compared two-level interventions with C-ADR and ACDF and reported SF-36 PCS and MCS scores (0-100 scale). Success for these component scores was defined as postoperative score improvement of ≥15 points from baseline scores. The likelihood of improved function based on PCS success was similar for C-ADR and ACDF short term (2 RCTs, N=657, 76.5% vs. 69.3%, RR 1.11, 95% CI 0.88 to 1.46, I^2= 72.7%),[117,118] intermediate term (2 RCTs, N=639, 83.7% vs. 79.1%. RR 1.06, 95% CI 0.92 to 1.36, I^2=69.7%),[70,117] and long term (1 RCT, N=216, 76.4% vs. 71.0%, RR 1.08, 95% CI 0.91 vs. 1.27)[118] (**Figure 31**). The likelihood of MCS success was also similar for C-ADR and ACDF at short term (2 RCTs, N=657, 50.3% vs. 45.2%, RR 1.08, 95% CI 0.82 to 1.41, I^2= 43.9%),[117,118] intermediate term (2 RCTs, N=639, 62.3% vs. 65.3%, RR 0.98, 95% CI 0.85 to 1.18, I^2=0%)[70,117] and long term (1 RCT, N=216, 53.7% vs. 52.7%, RR 1.02, 95% CI 0.79 to 1.31)[118] (**Figure 32**).

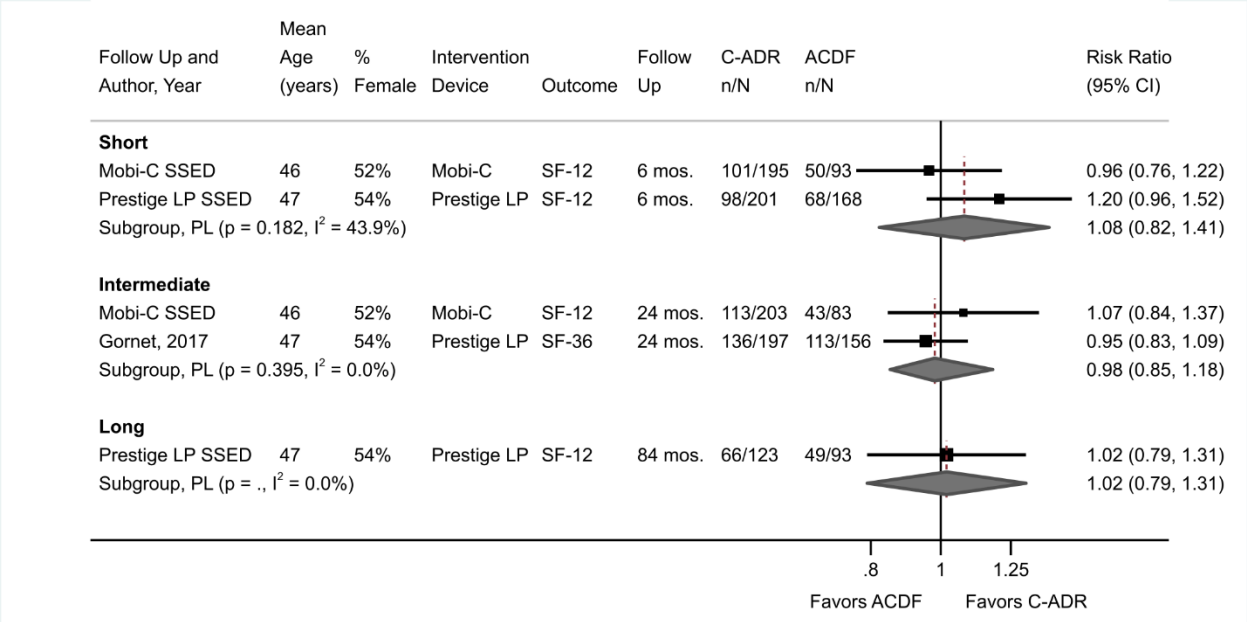
3. Results

Figure 31. SF-36 or SF-12 PCS success (≥15-point improvement): Comparison of C-ADR with ACDF (2-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PCS = Physical Component Score; PL = profile likelihood; SF-12 = Short-Form-12 questionnaire; SF-36 = Short-Form-36 questionnaire; SSED = Summary of Safety and Effectiveness Data (FDA).

Figure 32. SF-36 or SF-12 MCS success (≥15-point improvement): Comparison of C-ADR with ACDF (2-level interventions)

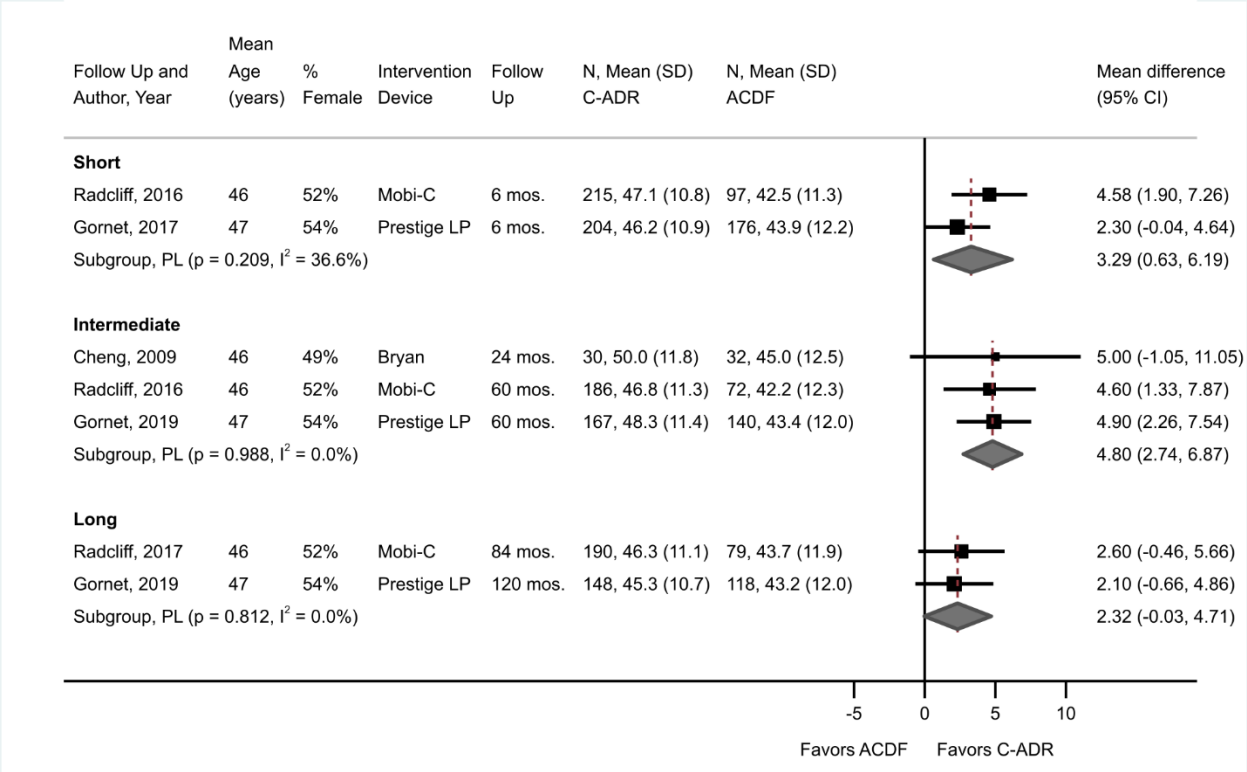


ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; MCS = Mental Component Score; mos. = months; PL = profile likelihood; SF-12= Short-Form-12 questionnaire; SF-36 = Short-Form-36 questionnaire; SSED = Summary of Safety and Effectiveness Data (FDA).

Three RCTs (N=792) (in 5 publications)[61,69,70,92,93] that compared two-level interventions with C-ADR and ACDF reported SF-36 PCS and MCS scores (0-100 scale). Differences in mean PCS scores did not meet the threshold for a small improvement and were classified as no

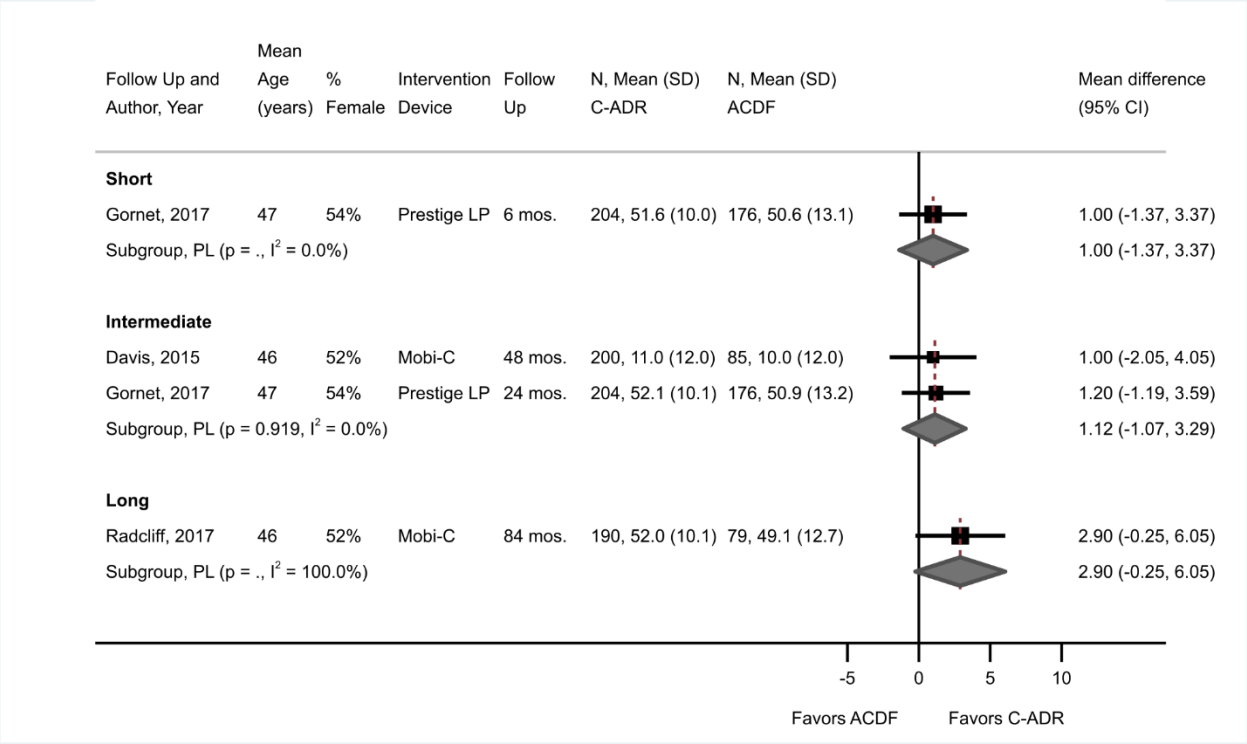
difference between C-ADR versus ACDF at short-term (2 RCTs, N=692, MD 3.29, 95% CI 0.63 to 6.19, I^2=36.6%),[70,92] intermediate term (3 RCTs, N=627, MD 4.80, 95% CI 2.74 to 6.87, I^2=0%)[61,69,92] and long-term (2 RCTs, N=535, MD 2.32, 95% CI -0.03 to 4.71, I^2=0%);[69,93] however, estimates were imprecise (**Figure 33**). Two RCTs (N=757) reported mean MCS scores there were also not different at short term (1 RCT, N=380, MD 1.00, 95% CI -1.37 to 3.37),[70] intermediate term (2 RCTs, N=665, MD 1.12, 95% CI -1.07 to 3.29, I^2=0%)[64,70] or long term (1 RCT, N=269, MD 2.90, 95% CI -0.25 to 6.05)[93] (**Figure 34**). One IDE NRSI (N=352) that compared a novel C-ADR versus matched historical ACDF controls found no difference in mean SF-36 PCS at short (49.2 vs. 46.4, p<0.05) or intermediate term (49.2 vs. 47.9).[102]

Figure 33. SF-36 or SF-12 PCS scores (0-100 scale): Comparison of C-ADR with ACDF (2-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PCS = Physical Component Score; PL = profile likelihood; SD = standard deviation; SF-12= Short-Form-12 questionnaire; SF-36 = Short-Form-36 questionnaire.

Figure 34. SF-36 or SF-12 MCS scores (0-100 scale): Comparison of C-ADR with ACDF (2-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; MCS = Mental Component Score; mos. = months; PL = profile likelihood; SD = standard deviation; SF-12= Short-Form-12 questionnaire; SF-36 = Short-Form-36 questionnaire.

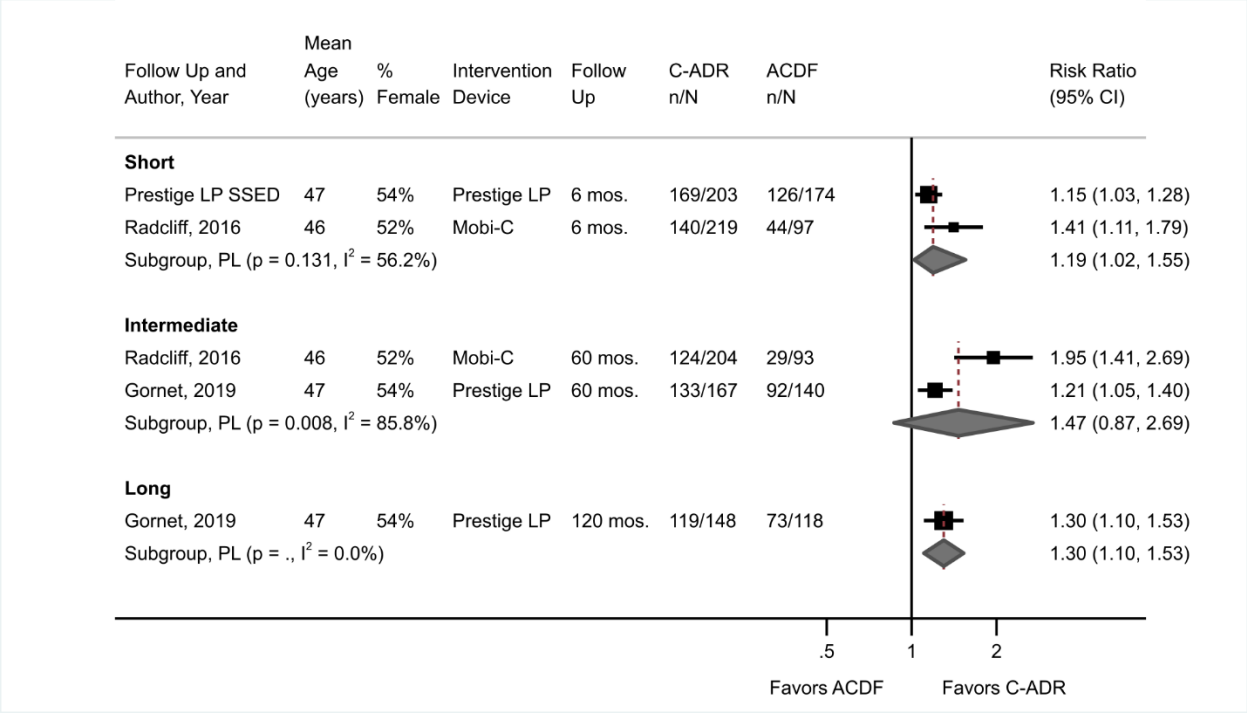
3.9.3.2.3.2.3 Odom's Criteria

There was no difference between C-ADR and ACDF for the likelihood of scoring excellent or good on Odom's criteria at intermediate term in one RCT (N= 62, 96.7% vs. 84.4%, RR 1.15, 95% CI 0.97 to 1.34).[61]

3.9.3.2.4 Overall Success (Composite)

The FDA IDE trials were required to report on overall success, a composite outcome that included a threshold of ≥15-point NDI improvement from baseline, improvement or maintenance of neurologic status, no serious adverse events and no additional surgical procedures that might be considered "failure" (e.g., removal, revision, supplemental fixation). C-ADR was associated with a slightly higher likelihood of overall success short term (2 RCTs, N=693, 73.2% vs. 62.7%, RR 1.19, 95% CI 1.02 to 1.56, I^2=56.2%) and long-term (1 RCT, N=266, 80.4% vs. 61.9%, RR 1.30, 95%CI 1.10to 1.53). At intermediate term, C-ADR was also associated with slightly greater likelihood of overall success in two RCTs individually (1 RCT N= 297, 60.1% vs. 31.2%, RR 1.95, 95% CI 1.41 to 2.69 and 1 RCT N=307, RR 1.12, 95%CI 1.05 to 1.40) (**Figure 35**).

Figure 35. Overall success (composite): Comparison of C-ADR with ACDF (2-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PL = profile likelihood; SSED = Summary of Safety and Effectiveness Data (FDA).

One IDE RCT defined overall success with different NDI success criteria (improvement from baseline of ≥30-points if baseline score was ≥60 or ≥50% if baseline score was <60), required adjudication of adverse events and added radiographic success to the criteria listed for the other IDE trials. C-ADR was associated with slightly higher likelihood of overall success long-term versus ACDF (1 RCT, N= 249, 60.8% vs. 34.6%, RR 1.76, 95% CI 1.27 to 2.44).[93] One IDE NRSI[102] that compared a novel C-ADR versus historical ACDF controls defined overall success as ≥15-point NDI improvement, maintenance or improvement in neurological status), no serious adverse event (any implant-associated or implant/surgical procedure–associated) and no additional index-level surgical procedure. Authors reported that overall success was more common in C-ADR participants versus ACDF (N=352, 86.7% vs. 77.1, p<0.05) based on multiple imputation modeling (numerators not reported; effect estimate could not be calculated).

3.9.3.2.5 Quality of Life

None of the included studies reported quality-of-life measures.

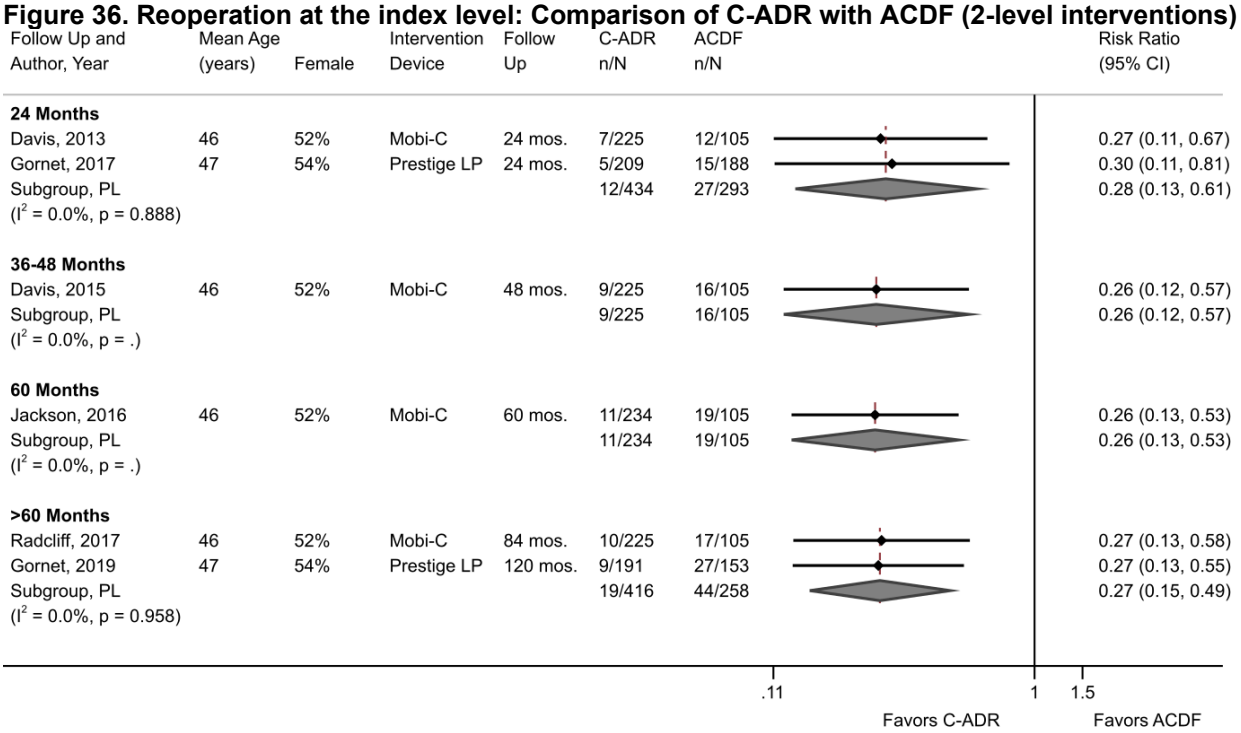
3.9.3.2.6 Reoperation

There was low-strength evidence that reoperation is substantially less likely with C-ADR compared with ACDF at all time points from 24 months and beyond (SOE: Low).

Reoperation included any additional procedure at the index level and was substantially less likely with C-ADR at all times reported across IDE trials, however estimates were imprecise. Effect estimates were consistent across reported times: up to 24 months (2 RCTs, N=727, 2.8% vs. 9.2%, RR 0.28, 95% CI 0.13 to 0.61, I^2=0%),[63,70] 36 to 48 months (1 RCT, N=330, 4.0% vs. 15.2%, RR 0.26, 95% CI 0.12 to 0.57),[64] 60 months (1 RCT, N=330, 4.7% vs. 18.1%, RR 0.26, 95% CI 0.13 to 0.53)[78] and >60 months (2 RCTs, N=674, 4.6% vs. 17.1%, RR 0.27, 95% CI 0.15

to 0.49, I^2=0%)[69,93] (**Figure 36**). One IDE NRSI that compared a novel C-ADR versus historical ACDF controls also reported that secondary surgical interventions were less common with C-ADR (N=352, 2.2% vs. 8.8%).[102]

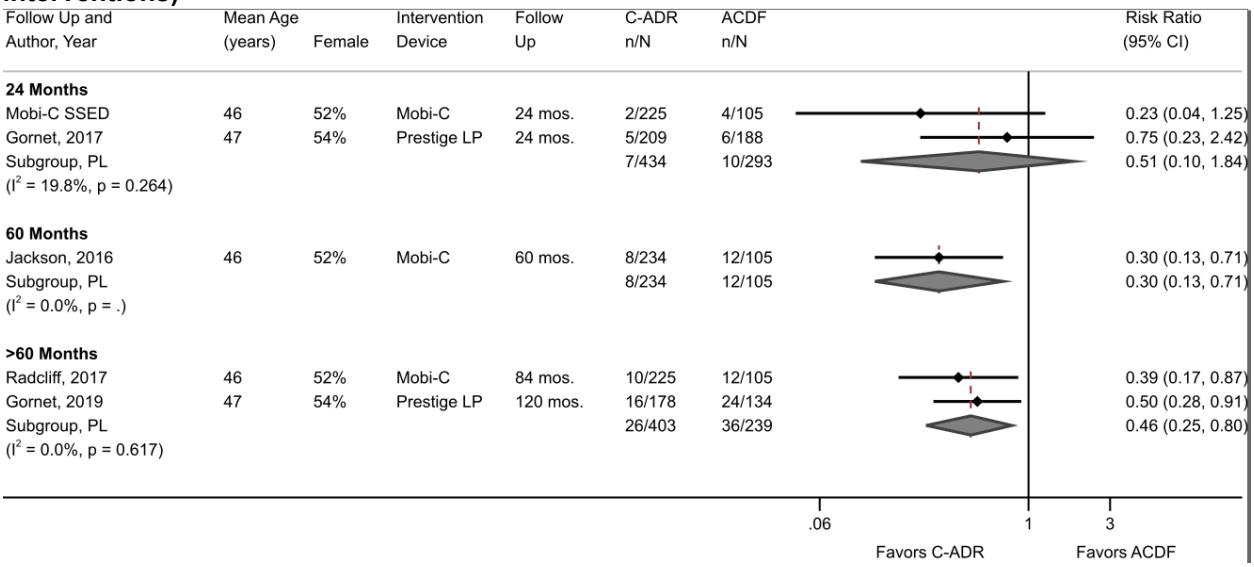
Figure 36. Reoperation at the index level: Comparison of C-ADR with ACDF (2-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PL = profile likelihood.

Subsequent surgery rates at adjacent levels were similar between C-ADR and ACDF at 24 months (2 RCTs, N= 727, 1.6% vs. 3.4%, RR 0.51, 95% CI 0.10 to 1.84, I^2=19.8%),[70,117] but substantially less common with C-ADR versus ACDF at 60 months (1 RCT, N=339, 3.4% vs. 11.4%, RR 0.30, 95% CI 0.13 to 0.71)[78] and >60 months (2 RCTs, N=642, 6.5% vs. 15.1%, RR 0.46, 95% CI 0.25 to 0.80, I^2= 0%).[69,93] Across trials, indications for operation at adjacent levels were not consistently described (**Figure 37**).

Figure 37. Subsequent surgery at adjacent level: Comparison of C-ADR with ACDF (2-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; mos. = months; PL = profile likelihood.

3.9.3.2.7 Harms

C-ADR was associated with a slightly lower likelihood of experiencing any adverse event at 24 months based on low-strength evidence (SOE: Low), but there was no difference between procedures at 120 months for WHO Grade 3 or 4 adverse events (SOE: Low). There was insufficient evidence for neurological deficits or events and for mortality (SOE: Insufficient).

All IDE RCTs and one IDE NRSI provided information on adverse events and harms.

3.9.3.2.7.1 Neurologic Deficit

Two RCTs (N=395) in 3 publications[61,64,93] reported neurologic events using varied terminology. One RCT (N=65)[61] reported that no neurologic complications occurred with C-ADR or ACDF through 24 months. There was no difference between neurologic deterioration at 48 months (6.2% vs. 7.6%, RR 0.82, 95% CI 0.35 to 1.89) in one IDE trial[64] but a subsequent publication of the trial reported substantially lower incidence of neurological failure, defined as a decrease in sensory, reflex or motor function from preoperative status, with C-ADR versus ACDF (6.4% vs. 17.1%, RR 0.36, 95% CI 0.19 to 0.70) at 84 months.[93]

3.9.3.2.7.2 Mortality

Cumulative mortality was similar between two-level C-ADR (2 deaths) and ACDF (3 deaths) through 120 months in one IDE trial, but authors did not provide cause of death (N=397, 1.0% vs. 1.6%; RR 0.60, 95% CI 0.10 to 3.55);[69] there was one death in both groups by 12 months (0.5% vs. 0.5%)[70] and two deaths in both groups by 84 months (1.0% vs. 1.1%).[81]

3.9.3.2.7.3 Serious Adverse Events

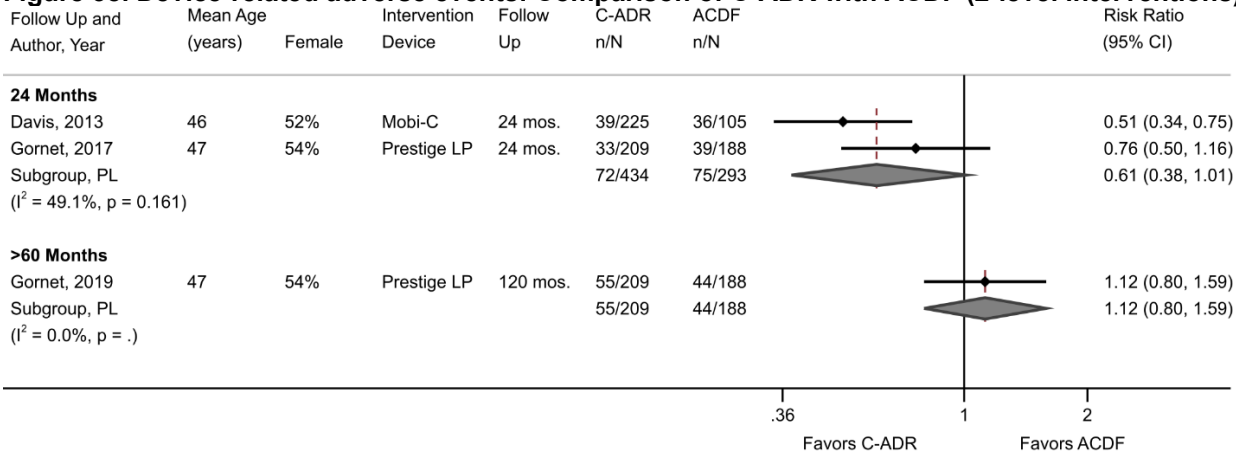
Serious adverse events were reported for two IDE trials (N=727) of different devices (five publications)[63,64,69,70,81] but were defined differently across reports. One trial's initial report found events were common and that fewer C-ADR (Mobi-C) participants experienced one or more serious adverse events (23.9% vs. 32.4%)[63] up to 24 months but included events unrelated to the device, surgery, or cervical spine as well as those that may not have required additional medical

intervention. In a subsequent report of this trial, following adjudication of events by a clinical events committee, fewer events were considered serious and they continued to be less common with C-ADR versus ACDF, but effect estimates were imprecise (1 RCT, N=330, 4.0% vs. 7.6%, RR 0.75, 95% CI 0.53 to 1.08) at 24 months.[64] The IDE trial of another device (Prestige-LP), also included a broad range of events and reported fewer Grade 3 or 4 adverse events with C-ADR at 24 months vs. ACDF (1 RCT, N=397, 34.4 % vs. 47.9%).[70] C-ADR was associated with slightly lower likelihood of serious AEs across the two trials at 24 months (2 RCTs, N=727, 29.3% vs. 42.3%, RR 0.73, 95% CI 0.58 to 0.93, I^2=0%)[63,70] using the broad definition of events. There was no difference between groups in the frequency of WHO Grade 3 or 4 adverse events at 120 months in one IDE trial (N=397, 66.7% vs. 70.9%, RR 0.93, 95% CI 0.80 to 1.09).[69]

3.9.3.2.7.4 Device-related Adverse Events

Device-related adverse event definitions, types of events and adjudication varied across RCTs. One trial included a range of events such as anatomy/technical difficulty, trauma as well as neurological events while others did not provide specifics. Some device-related events may only occur with C-ADR, others may only occur with ACDF (e.g., nonunion). Some events may not be persistent or serious (e.g., dysphagia or dysphonia). Two-level C-ADR was associated with a moderately lower likelihood of device-related events at 24 months compared with ACDF (2 RCTs, N=727, 16.6% vs. 25.6%, RR 0.61, 95% CI 0.38 to 1.01, I^2=49.1%)[63,70] but there was no difference between groups at 120 months in one of these trials (N=397, 26.3% vs. 23.4%, RR 1.12, 95% CI 0.80 to 1.59)[69] (**Figure 38**). When only serious device-related adverse events were considered, as adjudicated by committee or as WHO grade 3 or 4 events, C-ADR was associated with a substantially lower likelihood of such serious events compared with ACDF at 24 months in one trial (N=397, 1.9% vs. 5.9%, RR 0.33, 95% CI 0.11 to 1.01)[70] but there was no difference between groups at 120 months in this same trial (RR 0.45, 3.3% vs. 7.4%, 95% CI 0.19 to 1.09)[69] or at 60 months in a second trial (N=330, 4.4% vs. 8.6%, RR 0.52, 95% CI 0.22 to 1.24)[92] however, the estimates were very imprecise.

Figure 38. Device-related adverse events: Comparison of C-ADR with ACDF (2-level interventions)



ACDF = anterior cervical discectomy and fusion; C-ADR = cervical artificial disc replacement; CI = confidence interval; F/U = followup; mos. = months; PL = profile likelihood.

Device related AES were similar for C-ADR and ACDF in one IDE NRSI (3.8% vs. 3.5%).[102]

3.9.3.2.7.5 Dysphagia

Dysphagia was reported by several RCT publications (N=475), but the severity was unclear in most cases.[61,92,97] Dysphagia rate ranges were broad for C-ADR (0% to 24%) and for ACDF (0% to 38%) across these publications. One IDE trial (N=397) reported low rates of Grade 3 or 4 dysphagia that differed slightly across two post-FDA approval study publications, possibly reflecting different analytic methods. Rates did not differ by procedure at 84 months (1.3% vs. 0%)[81] or 120 months (0.6% vs. 0.7%).[69]

3.9.3.2.7.6 Heterotopic Ossification

Two IDE RCTs (N=434, C-ADR arms) reported heterotopic ossification (HO) for 2-level interventions. One trial[92] reported Grade 4 HO in 9.7 percent of C-ADR participants by 60 months (22/255) (not reported for ACDF group) while another reported no Grade 3 or 4 HO following C-ADR (N=209) versus three with ACDF (1.6%, 3/188) by 120 months.[69]

3.9.3.2.8 Differential Effectiveness (HTE)

One IDE trial that compared 2-level C-ADR and ACDF provided subgroup analysis on the presence of radiculopathy alone (N=287) and myelopathy alone or myelopathy with radiculopathy (N=110) for pain, function. and adverse events at 24 and 84 months but did not formally test for interaction.[71] Visual inspection of effect estimates and overlap in estimate variability and subgroup estimates suggest no differential effectiveness or harms, although the study may have been underpowered to evaluate this.

3.9.3.3 Mixed 1-, 2- or 3-level C-ADR versus ACDF

Three RCTs compared 1- 2- or 3-level C-ADR and ACDF (i.e., mixed levels).[60,62,72] Sample sizes ranged from 53 to 83 (total N=196). Across two trials,[60,62] 54 to 83 percent of participants had single-level procedures, 17 to 37 percent had 2-level procedures, and in one of these trials[60] 8 percent had 3-level procedures; one trial used the Bryan® disc and the other used the Prestige-II® disc, which are both FDA approved for single-level indications only. The third trial enrolled participants who underwent 1- or 2-level procedures but did not provide the proportions for each.[72] The RCTs were conducted in China, India and Spain. Two additional NRSIs compared harms for mixed-level C-ADR and ACDF.[105,108]

3.9.3.3.1 Fusion

One RCT (N=42) reported intermediate-term fusion success in 90.5 percent of participants in the ACDF arm.[60] This RCT also reported fusion in the C-ADR arm, but this can be attributed to participant crossover after initial randomization.

3.9.3.3.1 Pain

There was low-strength evidence of no difference between treatment with C-ADR and ACDF on neck pain (SOE: Low).

There was no difference in median VAS (0 to 10) neck pain scores at 60 months between C-ADR (3.6, interquartile range [IQR] 3.2 to 4.1) and ACDF (median 3.9, IQR 3.0 to 4.4) at 60 months (p=0.203) in one trial (N=50).[72] No other pain measures were reported.

3.9.3.3.3 Function

3.9.3.3.3.1 Neurologic Function

There was inadequate evidence to determine the effect of C-ADR versus ACDF on neurologic function. (SOE: Insufficient).

Participants who received C-ADR had higher mean JOA scores (0-17) at 36 months compared with ACDF in one RCT (N=81): 15.4 versus 14.7 (estimated from graphs in article), p=0.016.[60]

3.9.3.3.3.2 General Function

There was inadequate evidence to determine the effect of C-ADR versus ACDF on general function. (SOE: Insufficient).

One RCT (N=81) reported three different measures of general function at 36 months.[60] Participants who received C-ADR had better (i.e., lower) mean NDI scores (12 vs. 18 [estimated from graphs], on a 0 to 50 scale, p<0.001) and better (i.e., higher) mean SF-36 PCS scores (50.5 vs. 44.5 [estimated from graphs], on a 0 to 100 scale, p<0.05) compared with ACDF, but there were no differences between treatments in the proportion of participants who achieved an excellent (58.5% vs. 58.5%, RR 1.02, 95% CI 0.70 to 1.47) or good (34.1% vs. 25%, RR 1.37, 95% CI 0.69 to 2.71) result according to Odom's criteria. A second RCT (N=50) reported no difference between groups in NDI scores (median 7, IQR 6 to 8, for both groups) at 60 months.[72]

3.9.3.3.4 Quality of Life

None of the included studies reported on quality-of-life measures.

3.9.3.3.5 Harms

There was inadequate evidence to determine the effect of C-ADR and ACDF on harms or adverse events (SOE: Inadequate).

Two RCTs[60,62] and two NRSIs[105,108] reported harms and adverse events.
[60,62,108]

3.9.3.3.5.1 Neurological Complications

One RCT (N=53) reported one case of transient recurrent nerve paralysis in both groups (C-ADR 4% vs. ACDF 3.6%, RR 1.12, 95% CI 0.07 to 16.98) that resolved within 3-4 weeks and one case of postoperative worsening of arm pain and neurological deficit in the ACDF group (3.6%).[62] A second trial (N=83) reported that no intraoperative neurologic complications occurred in either group.[60] One large NRSI based on administrative data reported no difference between C-ADR and ACDF in the frequency of neurological complications (C-ADR 1.6% vs. ACDF 1.7%, adjusted OR 1.18, 95% CI 0.38 to 3.72), however specific types or timing of neurological events were not reported.[105]

3.9.3.3.5.2 Mortality

One RCT (N=83) reported that no deaths occurred in either group through 90 months.[60] Mortality was rare for both C-ADR (0.5%) and ACDF (2.2%) and there was no difference between procedures (OR 0.56, 95% CI 0.08 to 4.11) in one large NRSI based on administrative data.[105]

3.9.3.3.5.3 Serious Adverse Events

One RCT (N=83) reported one case of DVT (2.4%) in the C-ADR group.[60] There were no differences between C-ADR and ACDF in the frequency of pulmonary embolism (0.5% vs. 0.8%, OR 1.43, 95% CI 0.19 to 10.7) or deep vein thrombosis (2.2% vs. 2.4%, OR 1.07, 95% CI 0.33 to 3.40) in one large NRSI (N=143,060).[105]

One RCT (N=83) reported that no cerebrospinal fluid (CSF) leakage occurred.[60] CSF leak was rare for both C-ADR (0.5%) and ACDF (0.2%) and there was no difference between procedures (OR 2.19, 95% CI 0.29 to 16.3) in one large NRSI based on administrative data.[105]

In one RCT (N=53), one participant (3.6%) who underwent 2-level ACDF developed a wound hematoma that needed urgent evacuation;[62] another RCT reported that there were no cases of wound hematoma.[60] One of these trials reported that three ACDF participants (10.7%, N=28) had recurrent cervical pain between 3 and 6 months which required local infiltration (not further explained).[62]

One case (2.4%, N=41) of heterotopic ossification was reported in the C-ADR group in another RCT.[60]

Although dysphagia was reported in one RCT[60] and one NRSI,[105] the severity of dysphagia was unclear.

3.9.3.3.6 Reoperation and Subsequent Surgery

One RCT (N=53) reported reoperation at the index level in one (4%) C-ADR and two (7.1%) ACDF participants between 12 and 36 months (RR 0.56, 95% CI 0.05 to 5.81).[62] A second trial (N=83) reported that no participants in either group required reoperation at the index level through 36 months.[60] One NRSI did not provide adjusted effect estimates but reported the proportions of C-ADR and ACDF patients who required reoperation at the index level at 12 months (1.7% vs. 2.4%) and 24 months (0% vs. 3.6%) and subsequent surgery at adjacent levels at 12 months (1.7% vs. 2.4%) and 24 months (3.3% vs. 5.1%).[108]

3.9.3.3.7 Differential effectiveness (HTE)

None of the included trials that compared 1-, 2-, or 3 level C-ADR and ACDF interventions reported differential effectiveness based on patient or other characteristics.

3.10 Key Question 9. In patients undergoing anterior cervical discectomy and fusion, what are the comparative effectiveness and harms of surgery based on interbody graft material or device type?

3.10.1 Standalone Device Versus Traditional Plate and Cage

3.10.1.1 Key Findings

- There was moderate-strength evidence of no difference in fusion rates between standalone cages versus plate and cage (SOE: Moderate).
- There was low-strength evidence of no differences between standalone cages versus plate and cage on improvement in arm pain, function, and quality of life (SOE: Low); there was inadequate evidence for neck pain improvement (SOE: Insufficient).
- There was low-strength evidence of no difference between standalone cage versus plate and cage on adjacent-level ossification (SOE: Low); evidence was inadequate for subsidence (sinking of vertebral endplates around the graft) and other adverse events (SOE: Insufficient).

3.10.1.2 Description of Included Studies

Nine RCTs (N=619)[120-128] compared a standalone device with a traditional plate and cage (**Appendix C**). The average mean followup duration was 21 months (range immediately postoperative to 36 months). Six trials were conducted in China, two in the U.S., and one each in Germany and Japan.

The average study mean age of participants was 52 years (range 41 years to 63 years); the average proportion of females was 42% (range 9% to 54%). Few trials reported exact proportions of patients with radiculopathy, myelopathy, or myeloradiculopathy. One trial enrolled only participants with radiculopathy without myelopathy[126] and two trials enrolled only participants with myelopathy but did not report the proportion of participants with radiculopathy.[123,125] Most trials enrolled participants with 1-level disease,[122,124,126] 1- to 2-level disease,[127,128] or 2-level disease.[121] One trial each treated participants with 1- to 3-level disease,[120] 3-level disease,[123] and 2- to 4-level disease.[125]

All studies were rated moderate risk of bias with the exception of one trial that was rated high risk of bias (**Appendix D**).[122] Methodological limitations included unclear randomization techniques, unclear blinding, and unclear attrition. Evidence for neck pain in standalone devices versus traditional plate and cage was rated insufficient due to conflicting findings. Evidence for harms other than adjacent-level ossification was rated insufficient due to the infrequency of adverse events (**Appendix G**).

3.10.1.3 Detailed Analysis

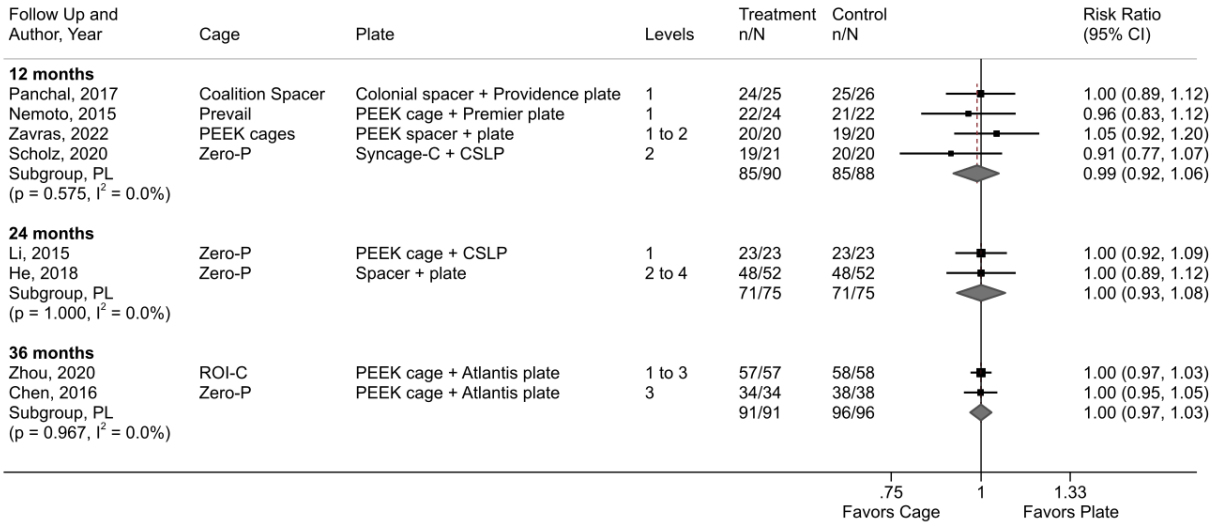
3.10.1.3.1 Fusion

There was moderate-strength evidence of no difference in fusion rates between standalone cages versus plate and cage in participants undergoing ACDF (SOE: Moderate).

Almost all participants who underwent ACDF with either a standalone cage or with a traditional plate and cage (N=515) experienced fusion at 12 months (4 RCTs, N=178, 94% vs.

97%, RR 0.99, 95% CI 0.92 to 1.06, I^2=0%), 24 months (2 RCTs, N=150, 95% vs. 95%, RR 1.00, 95% CI 0.93 to 1.08, I^2=0%) and 36 months (2 RCTs, N=187, 100% vs. 100%, RR 1.00, 95% CI 0.97 to 1.03, I^2=0%) (**Figure 39**). This was true when fusion was limited to one level or involved multilevel fusion. One trial did not report fusion as an outcome.[127]

Figure 39. Fusion, standalone cage vs. traditional plate and cage



CSLP = cervical spine locking plate; CI = confidence interval; PEEK = polyetheretherketone; PL = profile likelihood; ROI-C = ROI-C implant system; Zero-P = zero-profile

3.10.1.3.2 Pain

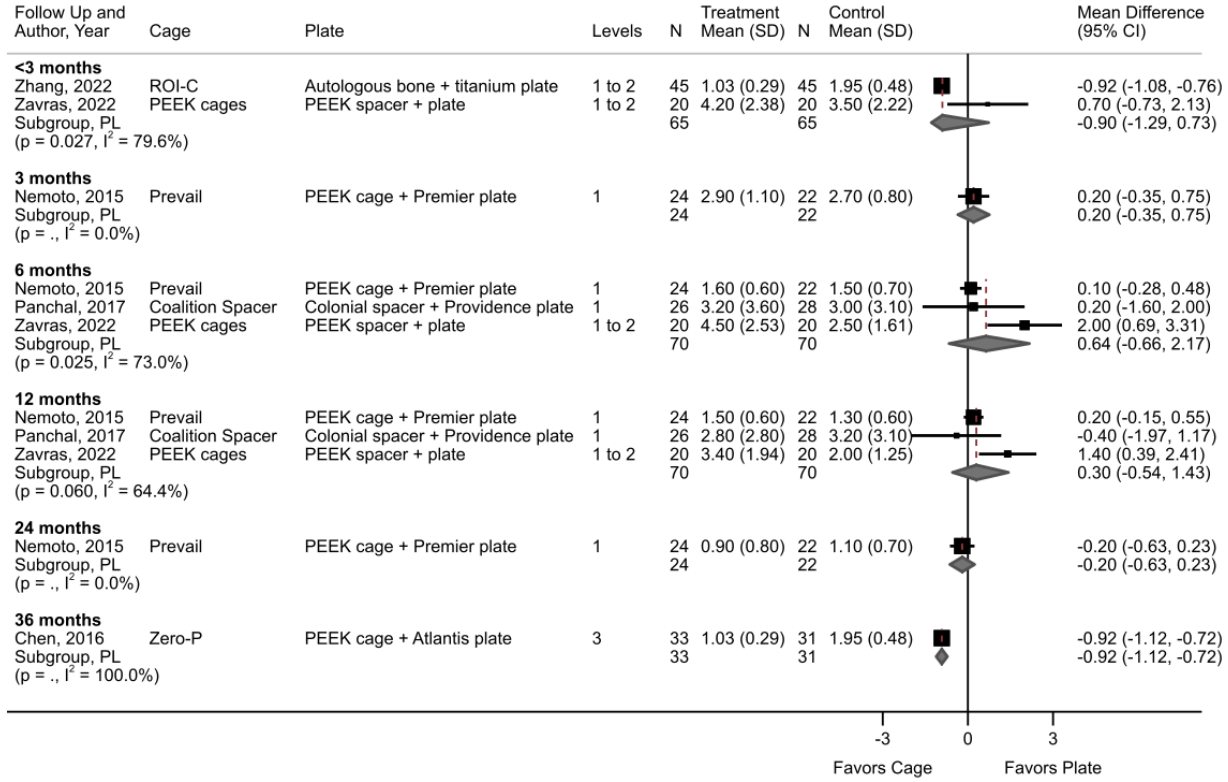
There was low-strength evidence of no difference between standalone cages versus plate and cage on improvement in arm pain (SOE: Low), with inadequate evidence to determine the benefits and harms of the two approaches on neck pain (SOE: Insufficient).

Five RCTs (N=294) reported changes in overall pain (pain location not specified) or neck pain using a visual analogue scale (VAS: 0-10 or 0-100) across various followup times ranging from less than 3 months to 36 months (**Figure 40**). Although neck pain improvement was moderately greater at less than 3 months and at 36 months, and statistically significant at 36 months, with a standalone cage compared with plate and cage, the opposite was true at 6 months (moderately, but not statistically more improved with plate and cage). When pooled analysis was limited to trials of single-level disease, there were no differences in neck pain between standalone cage and plate and cage (**Appendix F, Figure 3**).

Four RCTs (N=186) reported changes in arm pain using a visual analogue scale (VAS: 0-10 or 0-100) across various followup times. There were no differences in arm pain improvement after ACDF between use of a standalone cage and a plate and cage (**Figure 41**). When analyses were limited to trials of single-level disease, there remained no difference in improvement in arm pain between fusion methods (**Appendix F, Figure 4**).

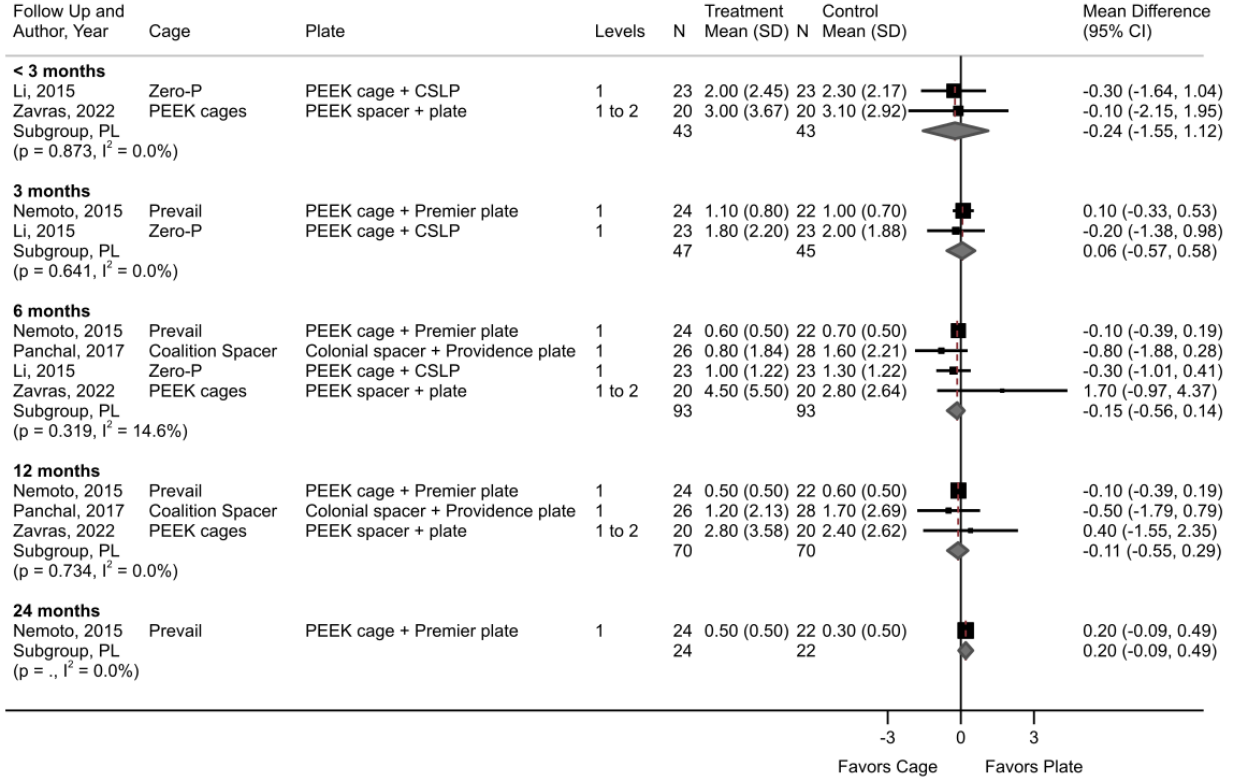
3. Results

Figure 40. Improvement in neck/unspecified pain after ACDF



ACDF = anterior cervical discectomy and fusion; CI = confidence interval; PEEK = polyetheretherketone; PL = profile likelihood; ROI-C = ROI-C implant system; SD = standard deviation; Zero-P = zero-profile

Figure 41. Improvement in arm pain following ACDF



ACDF = anterior cervical discectomy and fusion; CI = confidence interval; CSLP = cervical spine locking plate; PEEK = polyetheretherketone; PL = profile likelihood; ROI-C = ROI-C implant system; SD = standard deviation; Zero-P = zero-profile

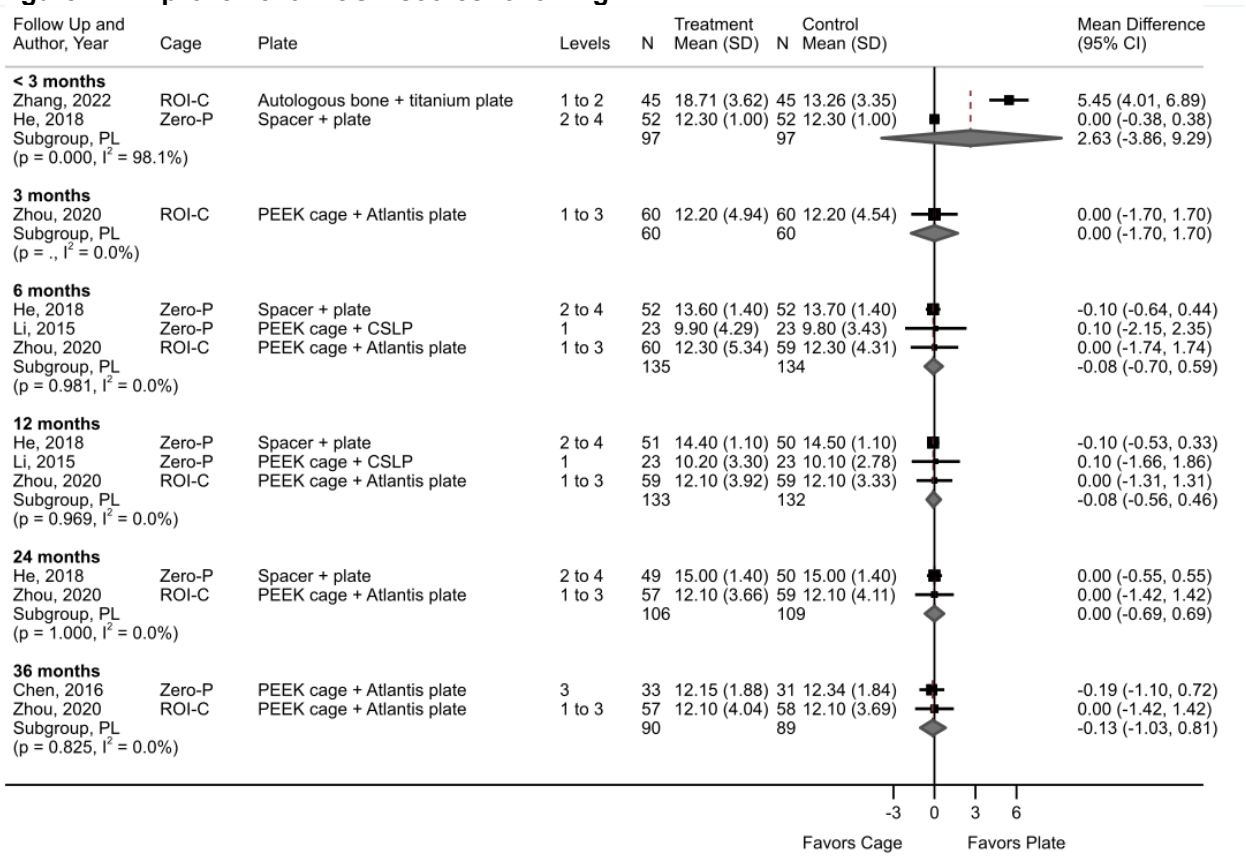
3.10.1.3.3 Function

3.10.1.3.3.1 Neurologic Function

There was low-strength evidence of no difference between standalone cages versus plate and cage in neurologic function (SOE: Low).

Five RCTs (N=424) reported changes on the Japanese Orthopedic Association score (JOA, lower score = worse disability, score 0 to 17) after ACDF using a standalone cage or a plate and cage (**Figure 42**). At less than 3 months, pooled analysis of two trials indicated a moderately greater, although not statistically significant, improvement in JOA scores with a standalone cage versus a plate and cage (MD 2.63, 95% CI -3.86 to 9.29), this effect is driven by 1 of 2 trials, while the other trial found no effect. At longer followup times, there were no differences between treatments on improvement in JOA scores.

Figure 42. Improvement in JOA scores following ACDF



ACDF = anterior cervical discectomy and fusion; CI = confidence interval; CSLP = cervical spine locking plate; PEEK = polyetheretherketone; PL = profile likelihood; ROI-C = ROI-C implant system; SD = standard deviation; Zero-P = zero-profile

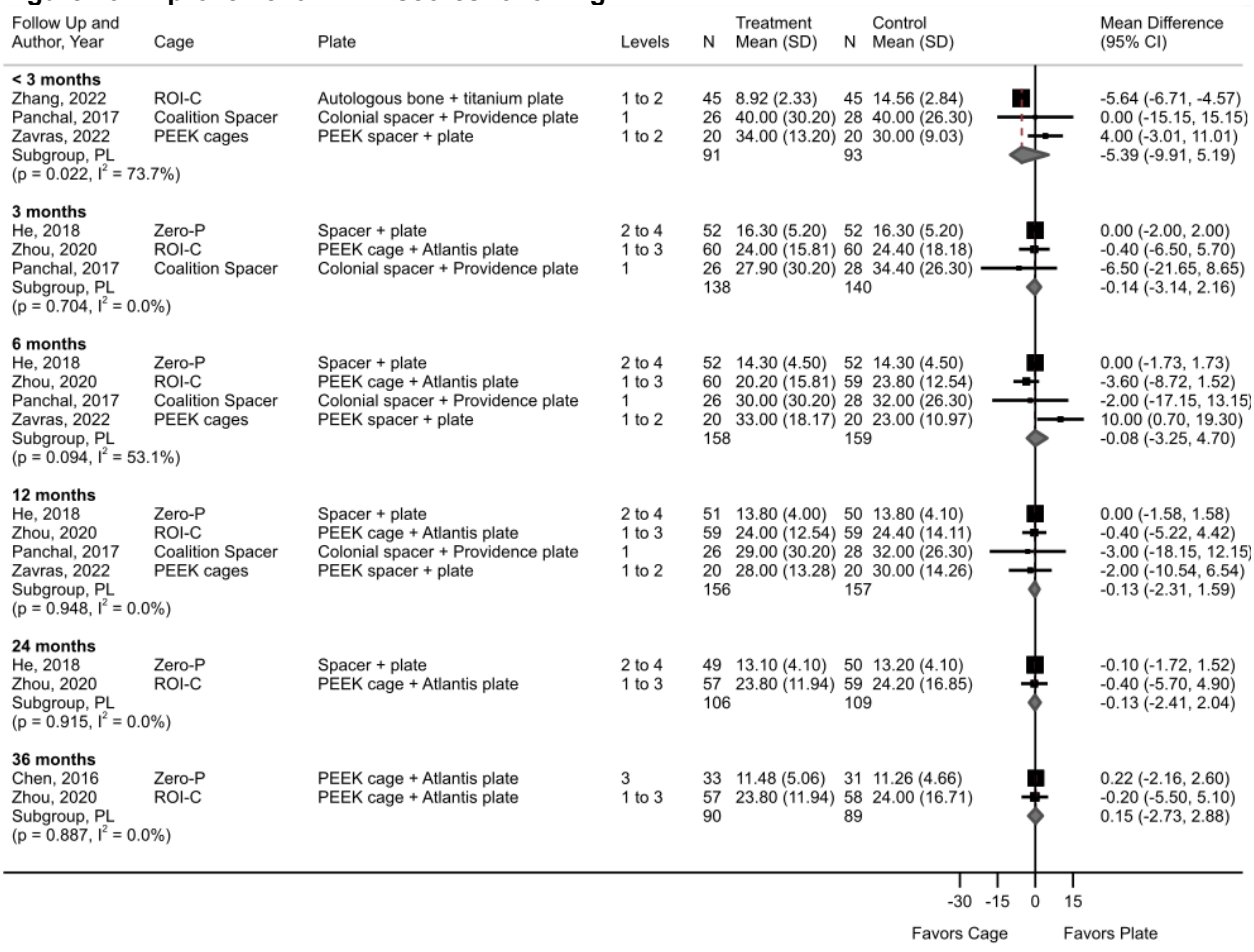
3.10.1.3.3.2 General Function

There was low-strength evidence of no difference between standalone cages versus plate and cage in general function (SOE: Low).

Six RCTs (N=472) reported changes on the Neck Disability Index (NDI, higher score = worse disability, 0-50 raw score or 0% to 100%) following ACDF with either a standalone cage

or a plate and cage (**Figure 43**). With the exception of less than 3 months timepoint, there were no differences between ACDF with a standalone cage or plate and cage on improvement in NDI scores. At less than 3 months, study findings varied and although the pooled estimate slightly favors the standalone cage, it is driven by the largest of the three studies and should interpreted with caution.

Figure 43. Improvement in NDI scores following ACDF



ACDF = anterior cervical discectomy and fusion; CI = confidence interval; NDI = Neck Disability Index; PEEK = polyetheretherketone; PL = profile likelihood; ROI-C = ROI-C implant system; SD = standard deviation; Zero-P = zero-profile

Additionally, one trial (N=41) reported no difference at 24 months between a standalone zero-profile device (Zero-P) and a plate and cage on the German version of the Neck Pain Disability Index (25.8% vs. 22.2%, p-value not reported).[121]

One RCT (N=46) reported no difference between a standalone cage and plate and cage at 24 months on the Odom's Criteria (Excellent: 46% vs. 55%; Good: 54% vs. 45%; Fair: 0% vs. 0%; Bad: 0% vs. 0%),[126] while another trial (N=41) reported the mean Odom's Grade at 24 months was 3.2 with a standalone cage compared with 3.5 with plate and cage (p-value not reported).[121] A third trial (N=115) reported there were no differences between standalone cage versus plate and cage in ratings of "excellent" and "good" overall patient satisfaction (Excellent: 44% vs. 47%, p=0.763; Good: 33% vs. 29%, p=0.835; Fair: 23% vs. 24%, p=0.692; Poor: 0% vs. 0%, p=1.0) at 36 months.[120]

3.10.1.3.4 Quality of Life

There was low-strength evidence of no difference between standalone cages versus plate and cage in quality of life (SOE: Low).

One RCT (N=40) reported no differences in quality of life as assessed with the Veteran's RAND 12-Item Health Survey between treatment with a standalone cage versus a plate and cage at 6 weeks and at 12 months, although participants treated with a standalone cage reported better scores at 6 months postoperatively (38.38 vs. 26.27, p=0.033).[128]

Five RCTs (N=253) assessed swallowing before and after treatment with a standalone cage versus a plate and cage with mixed results.[121-124,128] Two trials used the Swallowing Quality of Life (SWAL-QOL) questionnaire,[123,128] two trials rated severity of dysphagia symptoms as "None", "Mild", "Moderate", and "Severe"[121,124] and one trial used the Eating Assessment Tool.[122] No trial reported differences in dysphagia scores between treatments beyond 3 months postoperatively. One trial reported worse dysphagia scores with plate and cage immediately postoperatively, at 1 month, and at 3 months but no difference at 12 months.[124] Another trial reported worse scores with plate and cage at 6 weeks but no differences at 6 and 12 months.[128] There were no differences between dysphagia scores at any time from the postoperative period to 12 month in one RCT[122] and no differences at 36 months (only time reported) in another trial.[123] One trial reported no patient rated dysphagia as "moderate" or "severe" with either treatment[121] and no study reported that dysphagia required medical intervention (e.g., return to the operating room, PEG tube placement).

One RCT (N=54) rated high risk of bias found no differences on the Voice Handicap Index between treatment with a standalone cage versus plate and cage from discharge to 12 months.[122]

3.10.1.3.5 Harms

There was low-strength evidence of no difference between standalone cage versus plate and cage on adjacent-level ossification (SOE: Low), while evidence for subsidence and other adverse events was inadequate (SOE: Insufficient).

Seven RCTs (N=518) reported adverse events.[120,123-128]

Three trials reported substantially less adjacent-level ossification development (ALOD) with a standalone cage than with plate and cage (N=239, 8% vs. 27%, RR 0.25, 95% CI 0.12 to 0.52, I^2=8%). The change in ALOD severity grade (0=no ossification, 3=severe ossification) was reported in one study and favored treatment with the standalone cage (0.208 vs. 0.818, p=0.001).[126] (SOE: Low) However, no patient required reoperation at 36 months in two trials;[120,123] reoperation rates were not reported in the third trial.[126]

One RCT (N=46) reported a small, but not statistically significant difference in subsidence (loss of disc height) rates with a standalone cage compared with a plate and cage at 12 months (12.5% vs. 9.1%, RR 1.38, 95% CI 0.25 to 7.48) and at 24 months (16.7% vs. 13.6%, RR 1.22, 95% CI 0.31 to 4.87).[126]

One trial (N=104) reported few total complications (N=11) in 24 months that included 1 nerve injury (2%) and no cerebrospinal fluid leaks (0%) with the standalone cage compared with 2 nerve injuries (4%) and 1 cerebrospinal fluid leak (2%) with the plate and cage (p=0.999; p=1.00, respectively).[125] One trial (N=90) reported 1 (2%) incidence of loosening of the internally fixed implant with the standalone cage versus 3 (7%) with plate and cage (p=0.333).[127] Another trial (N=40) reported participant treated with a standalone cage experienced a screw loosening, interbody subsidence, and C-5 fracture with revision surgery under consideration at trial publication.[128] The same trial also reported one participant treated with a plate and cage

experienced screw fracture, pseudarthrosis and underwent posterior fusion and decompression 14 months after the primary surgery.

3.10.2 Titanium versus PEEK cages

3.10.2.1 Key Findings

- There was low-strength of greater likelihood of fusion with a PEEK cage compared with a titanium or titanium-coated PEEK cage (SOE: Low).
- There was low-strength evidence of greater likelihood of improved general function with a PEEK cage versus a titanium cage (SOE: Low); evidence for neurologic function was inadequate (SOE: Insufficient).
- Evidence for subsidence and other adverse events was inadequate (SOE: Insufficient).

3.10.2.2 Description of Included Studies

Three RCTs (N=217) compared ACDF using a titanium cage or titanium covered PEEK cage versus a PEEK cage.[129-131] (**Appendix C**) The average study mean duration of followup was 45 months (range 12 months to 99.7 months). One study each was conducted in China, Taiwan, and Poland.

The average study mean age of participants was 50 years (range 46 years to 52 years); the average proportion of female participants was 49% and 45% with one trial reporting that 72% of 170 disc spaces belonged to women. Two RCTs reported radiculopathy was experienced by 3% and 75%, myelopathy by 11% and 57%, and myeloradiculopathy by 13% and 40%.[129,130] The third trial did not report myeloradicular symptoms. One trial enrolled participants with 1-level (66%) or 2-level (34%) disease,[130 3803] 3-level disease[129] or disease at 1 or more levels[131]

All studies were rated moderate risk of bias (**Appendix D**). Methodological limitations included unclear randomization techniques, unclear blinding, and lack of intention to treat analysis. No funds were received in one trial[129] and funding was not reported in the other two. Evidence for neurologic function was rated insufficient due to limited evidence from one small trial. Evidence for subsidence was rated insufficient due to conflicting findings, while evidence for other harms was insufficient due to few adverse events (**Appendix G**).

3.10.2.3 Detailed Analysis

3.10.2.3.1 Fusion

There was low-strength evidence of a greater likelihood of fusion with a PEEK cage compared with a titanium or titanium-coated PEEK cage (SOE: Low)

Three RCTs (N=217) reported ACDF fusion rates at different followup times that were not different between titanium and PEEK cages or that favored PEEK cages.

One trial reported that at a mean of 99.7 months (range 86 to 116 months) all participants (N=60) achieved fusion of their 3-level disease with both the titanium cage and with the PEEK cage (87/87 levels vs. 93/93 levels).[129] However, followup was not available for 25% of the original participants. A second trial (N=53) reported a lower likelihood of fusion with the titanium cage (32/37 levels, 86.5%) versus the PEEK cages (34/34 levels, 100%, p=0.0335) after 24 months.[130] The third RCT (N=104) reported a large difference in the likelihood of complete fusion that favored the PEEK cage with complete fusion achieved in 26 of 59 titanium-covered

PEEK cages implanted (44.1%) compared with 75 of 85 PEEK cages implanted (88.2%) at 12 months ($p<0.001$).[131] Partial fusion was achieved by 55.9% of participants with titanium-covered PEEK cages and 11.76% of participants with PEEK cages.[131] There were no instances of an absence of fusion.[131]

3.10.2.3.2 Function

3.10.2.3.2.1 Neurologic Function

There was inadequate evidence of the benefits and harms of PEEK cage versus titanium cage on neurologic function (SOE: Insufficient).

One RCT (N=60) found JOA scores improved from baseline (baseline: 9.6 vs. 9.8) with both a titanium implant and a PEEK implant, but improvement was moderately greater with the PEEK implant (12.8 vs. 14.2, endpoint difference: -1.4, 95% CI -2.33 to -0.47).[129]

3.10.2.3.2.2 General Function

There was low-strength evidence of improved general function with a PEEK cage compared to a titanium cage (SOE: Low).

The same trial above (N=60) also found moderately improved NDI scores from baseline (baseline: 36.2 vs. 35.4) with both the titanium and the PEEK implant, but improvement was greater with the PEEK implant (21.6 vs. 15.2, endpoint difference: 6.4, 95% CI 5.13 to 7.67).[129]

Two RCTs (N=113) reported results on Odom's criteria that favored PEEK cages, although differences were not statistically significant in one trial.[129,130] One trial (N=60) reported moderately worse clinical status according to Odom's criteria with the titanium cage versus the PEEK cage (Excellent: 24% vs. 35%; Good: 31% vs. 39%; Fair: 28% vs. 16%; Bad: 17% vs. 10%, $p<0.05$).[129] One trial (N=53) reported no difference between treatments on clinical status (Excellent: 21% vs. 28%; Good: 54% vs. 52%; Fair: 14% vs. 8%; Poor: 11% vs. 12% or successful treatment: 75% vs. 80%, $p=0.6642$).[130] In the trial where enrollment was limited to individuals with 3-level disease, treatment with the PEEK cage was associated with better clinical status, whereas in the trial of 1- and 2-level disease, there was no differences between cage materials on perceived improvement. Additionally, the followup times were greatly different between trials (99.7 months vs. 24 months) with the longer followup time associated with better ratings.

3.10.2.3.3 Quality of Life

No studies reported quality of life outcomes.

3.10.2.3.4 Harms

Evidence was inadequate to determine the effect of a PEEK cage versus a titanium cage on subsidence or other adverse events (SOE: Insufficient).

One RCT (N=104) found no difference between a titanium-coated PEEK implant and a PEEK implant on the incidence of subsidence in 166 levels (20.6% vs. 21.4%, $p=0.875$).[131] However, subsidence was reported with 34.5% of titanium cages (87 levels) compared with 5.4% of PEEK cages (93 levels) in a second RCT (N=60, $p<0.05$)[129] and 16.2% of 37 levels versus 0% of 34 levels in a third RCT (N=53, $p<0.001$).[130] All three trials defined subsidence similarly (≥ 3 mm of interspace collapse). It is unclear the reason for the difference in study findings; possibilities include the cage materials (a titanium-coated PEEK cage may perform differently

than a titanium cage) and the duration since ACDF (12 months in the trial that found no difference versus 24 months and 99.7 months in the other two trials). (SOE: Insufficient)

One RCT (N=53) reported that after 24 months, there were no neurovascular injuries and no revision surgeries with either the titanium cage or the PEEK cage, but that one patient, who received the titanium cage, experienced a hematoma that was removed the day after surgery.[130] One RCT (N=60) reported that at a mean of 99.7 months two patients treated with a titanium cage experienced cage dislocation but were asymptomatic.[129]

3.10.3 Autograft, Allograft, and Other Osteogenic Materials

3.10.3.1 Key Findings

- There was inadequate evidence to determine comparative benefits (fusion, pain reduction, improved function, improved quality of life) for any osteogenic material versus any other osteogenic material (SOE: Insufficient).
- There was low-strength evidence that the use of BMP-2 in the cervical spine was associated with increased complications compared to no BMP-2 (SOE: Low); evidence was inadequate to determine the comparative harms of other osteogenic materials (SOE: Insufficient).

3.10.3.2 Description of Included Studies

Six RCTs (N=637) compared autologous bone graft, allograft, and/or other materials to support fusion in ACDF (**Appendix C**).[132-137] The average mean followup duration was 17 months (range 6 months to 24 months). Two trials were conducted in the U.S., two in China, and one each in South Korea and India.

The average study sample size was 106 (range 32 to 319); the average study mean age was 49 years (range 43 years to 55 years). One trial did not report age of participants.[135] The mean proportion of females enrolled was 52% (range 30% to 66%). The average proportion of patients with radiculopathy was 61% (range 28% to 100%); the average proportion of patients with myelopathy was 21% (range 0% to 38%), the average proportion of patients with myeloradiculopathy was 18% (range 0% to 34%). One trial reported that all study participants had radiculopathy, myelopathy or both.[136] All participants enrolled had 1-level degenerative disease,[133,137] 1- to 2-level disease[132,134,136] or 1- to 3-level disease.[135]

Additionally, two NRSI (N=944) assessed heterotopic ossification and complications due to neck swelling with the use of BMP-2 compared to anterior cervical fusion without BMP-2.[138,139] The mean age in one NRSI was 51 years with 51% female and 24% of study participants having myelopathy and 1 or more levels fused.[139] The other nonrandomized study, which took data from multiple investigational device exemption trials, did not report aggregate baseline patient characteristics but used propensity scoring on 28 predefined demographic and preoperative variables.[138]

One RCT was rated high risk of bias[135] and the remaining RCTs were rated moderate risk of bias (**Appendix D**). Methodological limitations included unclear randomization methods, unclear blinding, and unclear attrition. Both NRSIs were also rated moderate risk of bias and were downgraded due to baseline differences between study groups on prognostic variables and unclear blinding of outcome assessor. Two trials each reported industry funding, nonprofit funding, and grant funding; one trial did not address funding. One NRSI used data from three Investigational Device Exemption (IDE) trials,[138] while the other reported no funds or support

from industry.[139] Evidence comparing allograft, autograft, and other osteogenic materials on likelihood of fusion, pain improvement, function, and overall harms (with the exception of BMP-2 use) was rated insufficient due to limited evidence for each comparison (**Appendix G**).

3.10.3.3 Detailed Analysis

3.10.3.3.1 Fusion

There was inadequate evidence to determine the comparative benefits and harms of autograft, allograft, or other osteogenic material versus any other osteogenic material on fusion (SOE: Insufficient).

Six RCTs (N=534) assessed ACDF with autograft, allograft, or other materials (e.g., hydroxyapatite, calcium sulphate) and found no differences between materials in achievement of spinal fusion (**Table 3**). Fusion rates for all materials were high for all trials but only one randomized study was available for each comparison.

Table 3. Fusion with ACDF using various osteogenic materials

Trial (Timepoint)	Intervention A (Sample size)	Intervention B (Sample size)	Findings
Arnold, 2018 (24 months)	i-Factor + allograft ring (N=117)	Local graft + allograft ring (N=127)	97.30% vs. 94.44%, p=0.2513
Baskin, 2003 (24 months)	BMP-2 + allograft ring (N=10)	ICBG + allograft ring (N=10)	100% vs. 100%, p=1.0
Cho, 2005 (6 months)	Biphasic calcium phosphate ceramic + PEEK cage (N=50)	ICBG + PEEK cage (N=50)	100% vs. 100%, p=1.0
Kanna, 2021 (12 months)	Allograft + patient's blood + titanium cage (N=13)	Local graft + titanium cage (N=14)	100% vs. 100%, p=1.0 Fusion grade: (p=0.73) F: 23.2% vs. 28.6% F+: 38.4% vs. 42.8% F++: 38.4% vs. 28.6%
Xie, 2015 (12 months) (24 months)	Calcium sulphate + demineralized bone matrix + PEEK cage (N=34)	Autogenous iliac cancellous bone + PEEK cage (N=32)	12 month 104 levels, 24 month levels NR: 12 months: 94.3% vs. 100%, p=NR 24 months: 100% vs. 100%, p=1.0
Yi, 2015 (12 months)	Hydroxyapatite + demineralized bone matrix + PEEK cage (N=38)	B-tricalcium phosphate + hydroxyapatite + PEEK cage (N=39)	X-ray: 87% vs. 87%, p=1.0 CT: 87% vs. 72%, p=0.16

BMP-2 = bone morphogenetic protein; CT = computed tomography; ICBG = iliac crest bone graft; i-Factor = biologic bone graft made of a small peptide bound to an anorganic bone mineral; PEEK = polyetheretherketone

3.10.3.3.2 Pain

There was inadequate evidence to determine the comparative benefits and harms of autograft, allograft, or other osteogenic material versus any other osteogenic material on neck or arm pain (SOE: Insufficient).

Five RCTs (N=440) assessed neck and arm pain using a VAS or a numerical (pain) rating scale (**Table 4-5**). One small trial (N=27) reported a moderately greater decrease in neck pain 12 months after ACDF with a local graft and titanium cage than with allograft and titanium cage (MD -6.15 vs. -5.09, $p<0.05$).[132] Another trial (N=20) found a moderate, though not statistically significant, improvement in neck pain with BMP-2 and allograft ring versus iliac crest bone graft and an allograft ring on a 20-point numerical rating scale (MD 13.0 vs. MD 9.0, $p>0.05$).[136]

One trial (N=27) also found a substantially greater decrease in arm pain with local graft and a titanium cage compared with allograft and the same cage (MD -7.24 vs. MD -4.55, $p<0.05$)[132] (**Table 5**). However, these results should be interpreted with caution due to the trial's small

sample size. One RCT (N=26) reported a substantially greater reduction in arm pain at 24 months with BMP-2 and allograft ring compared with iliac crest bone graft and allograft ring on a 20-point numerical rating scale (-14 vs. -8.5, p<0.03).[136] However, as above, these results should be interpreted with caution due to the small sample size. One RCT (N=244) found that ACDF with i-Factor (bone graft made of a peptide bound to an inorganic bone mineral) and an allograft ring was associated with improved VAS arm pain scores at 24 months (1.56 v s. 1.95, p=0.0306) compared with local graft and an allograft ring.[133] However, this small difference in scores is below the threshold for a small effect and may not be clinically meaningful. One RCT (N=77) found a small, although not statistically significant, improvement in arm pain at 12 months with hydroxyapatite, demineralized bone matrix and a PEEK cage compared with β-tricalcium phosphate, hydroxyapatite and a PEEK cage (VAS: MD -4.2 vs. MD-3.6, p=0.27).[137]

There were no differences in neck or arm pain with other comparisons.

Table 4. Neck pain with ACDF using various osteogenic materials

Trial (Timepoint)	Intervention A (Sample size)	Intervention B (Sample size)	Findings
Arnold, 2018 (24 months)	i-Factor + allograft ring (N=117)	Local graft + allograft ring (N=127)	VAS endpoint: 1.79, 95% CI 1.33 to 2.24 vs. 2.25, 95% CI 1.78 to 2.72, p=0.4619
Baskin, 2003 (24 months)	BMP-2 + allograft ring (N=14)	ICBG + allograft ring (N=12)	20-point NRS: MD 13.0 vs. MD 9.0, p>0.05
Kanna, 2021 (12 months)	Allograft + patient's blood + titanium cage (N=13)	Local graft + titanium cage (N=14)	0-10 NPRS: MD -5.09 vs. MD -6.15, **p<0.05**
Xie, 2015 (24 months)	Calcium sulphate + demineralized bone matrix + PEEK cage (N=34)	Autogenous iliac cancellous bone + PEEK cage (N=32)	Improved VAS neck pain: 69% vs. 68%, p>0.05
Yi, 2015 (12 months)	Hydroxyapatite + demineralized bone matrix + PEEK cage (N=38)	B-tricalcium phosphate + hydroxyapatite + PEEK cage (N=39)	VAS: MD -1.6 vs. -1.8, p=0.82

BMP-2 = bone morphogenetic protein; ICBG = iliac crest bone graft; i-Factor = biologic bone graft made of a small peptide bound to an anorganic bone mineral; MD = mean difference; N(P)RS = Numeric Pain Rating scale; PEEK = polyetheretherketone; VAS = Visual Analogue Scale

Table 5. Arm pain with ACDF using various osteogenic materials

Trial (Timepoint)	Intervention A (Sample size)	Intervention B (Sample size)	Findings
Arnold, 2018 (24 months)	i-Factor + allograft ring (N=117)	Local graft + allograft ring (N=127)	VAS endpoint: 1.56, 95% CI 1.06 to 2.05 vs. 1.95, 95% CI 1.51 to 2.39, **p=0.0306**
Baskin, 2003 (24 months)	BMP-2 + allograft ring (N=14)	ICBG + allograft ring (N=12)	20-point NRS: MD -14.0 vs. -8.5, **p<0.03**
Kanna, 2021 (12 months)	Allograft + patient's blood + titanium cage (N=13)	Local graft + titanium cage (N=14)	0-10 NPRS: MD -4.55 vs. -7.24, **p<0.05**
Xie, 2015 (24 months)	Calcium sulphate + demineralized bone matrix + PEEK cage (N=34)	Autogenous iliac cancellous bone + PEEK cage (N=32)	Improved VAS arm pain: 70% vs. 68%, p>0.05
Yi, 2015 (12 months)	Hydroxyapatite + demineralized bone matrix + PEEK cage (N=38)	B-tricalcium phosphate + hydroxyapatite + PEEK cage (N=39)	VAS: MD –4.2 vs. -3.6, p=0.27

BMP-2 = bone morphogenetic protein; ICBG = iliac crest bone graft; i-Factor = biologic bone graft made of a small peptide bound to an anorganic bone mineral; MD = mean difference; N(P)RS = Numeric Pain Rating scale; PEEK = polyetheretherketone; VAS = Visual Analogue Scale

3.10.3.3.3 Function

3.10.3.3.3.1 Neurologic Function

There was inadequate evidence to determine the comparative benefits and harms of autograft, allograft, or other osteogenic material versus any other osteogenic material on neurologic function (SOE: Insufficient).

Three RCTs (N=192) reported changes in neurological status after ACDF (**Table 6**). One trial (N=100) found no differences between use of biphasic calcium phosphate ceramic plus a PEEK cage compared with iliac crest bone graft plus a peek cage on JOA score, or JOA recovery rate at 6 months post ACDF.[135] One trial (N=66) reported no difference between calcium sulphate plus demineralized bone matrix plus a PEEK cage versus autogenous iliac cancellous bone plus a PEEK cage in JOA scores at 24 months.[134] One trial reported neurologic success (i.e., maintenance or improvement in sensory and motor function) in all remaining participants at 24 months, while another trial reported that almost all participants (94.87% vs. 93.70%) experienced neurologic success, also at 24 months.[133]

Table 6. Neurologic function with ACDF using various osteogenic materials

Trial (Timepoint)	Intervention A (Sample size)	Intervention B (Sample size)	Findings
Arnold, 2018 (24 months)	i-Factor + allograft ring (N=117)	Local graft + allograft ring (N=127)	Neurologic success: 94.87% vs. 93.70%, p=0.6944
Baskin, 2003 (24 months)	BMP-2 + allograft ring (N=14)	ICBG + allograft ring (N=12)	Neurologic success: 100% vs. 100%, p=1.0
Cho, 2005 (6 months)	Biphasic calcium phosphate ceramic + PEEK cage (N=50)	ICBG + PEEK cage (N=50)	JOA score: MD 2.84 vs. 2.48, p=0.17 JOA recovery rate: 86.51% vs. 83.48%, p=0.22
Xie, 2015 (24 months)	Calcium sulphate + demineralized bone matrix + PEEK cage (N=34)	Autogenous iliac cancellous bone + PEEK cage (N=32)	JOA score: MD 3.62 vs. 3.22, p>0.05

BMP-2 = bone morphogenetic protein; ICBG = iliac crest bone graft; i-Factor = biologic bone graft made of a small peptide bound to an anorganic bone mineral; JOA = Japanese Orthopaedic Association; MD = mean difference; PEEK = polyetheretherketone

3.10.3.3.3.2 General Function

There was inadequate evidence to determine the comparative benefits and harms of autograft, allograft, or other osteogenic material versus any other osteogenic material on general function (SOE: Insufficient).

Four RCTs (N=374) assessed post ACDF neck disability with the Neck Disability Index (**Table 7**). One RCT (N=244) found that treatment with i-Factor plus an allograft ring in ACDF resulted in slightly, though not statistically significant, improvement on NDI endpoint scores at 24 months compared with local graft and allograft ring (22.33 vs. 25.66, p=0.5607).[133] One small trial (N=26) reported moderately greater improvement on the NDI after 24 months with BMP-2 and allograft ring compared with iliac crest bone graft and allograft ring (52.7 vs. 36.9, p<0.03).[136] Another small trial (N=27) reported moderately greater improvement on NDI scores after 12 months with local graft plus a titanium cage versus allograft plus titanium cage (MD 56.5 vs. MD 41.4, p<0.05).[132] There was no difference in improvement in NDI scores with hydroxyapatite/demineralize bone matrix plus PEEK cage versus β-tricalcium phosphate/hydroxyapatite plus PEEK cage at 12 months.[137]

Three RCTs (N=357) assessed general function using the SF-36 or the 2-item SF-12 (**Table 7**). Two trials found no difference in function on the SF-36 after ACDF using an allograft ring with either i-Factor or local graft[133] or using an allograft with either BMP-2 or an iliac crest bone

graft.[136] One small trial (N=27) reported moderately better function at 12 months using the 2-item SF-12 with local graft plus a titanium cage compared with the same cage and allograft infused with the participant's blood (MD 48.7 vs. 65.9, p<0.05).[132] However, care should be used in interpreting these results due to the small study sample size.

Table 7. General function with ACDF using various osteogenic materials

Trial (Timepoint)	Intervention A (Sample size)	Intervention B (Sample size)	Findings
Arnold, 2018 (24 months)	i-Factor + allograft ring (N=117)	Local graft + allograft ring (N=127)	NDI endpoint: 22.33, 95% CI 18.90 to 25.76 vs. 25.66, 95% CI 22.55 to 28.78, p=0.5607
Baskin, 2003 (24 months)	BMP-2 + allograft ring (N=14)	ICBG + allograft ring (N=12)	NDI improvement from preoperative scores: 52.7 vs. 36.9, p<0.03
Kanna, 2021 (12 months)	Allograft + patient's blood + titanium cage (N=13)	Local graft + titanium cage (N=14)	NDI: MD 41.4 vs. MD 56.5, p<0.05
Yi, 2015 (12 months)	Hydroxyapatite + demineralized bone matrix + PEEK cage (N=38)	B-tricalcium phosphate + hydroxyapatite + PEEK cage (N=39)	NDI: MD 22 vs. MD 20, p=0.62
Arnold, 2018 (24 months)	i-Factor + allograft ring (N=117)	Local graft + allograft ring (N=127)	SF-36 PCS endpoint: 45.40, 95% CI 43.60 to 47.20 vs. 44.47, 95% CI 42.70 to 46.24, p=0.6461 SF-36 MCS endpoint: 48.43, 95% CI 46.43 to 50.44 vs. 48.41, 95% CI 46.42 to 50.40, p=0.9040
Baskin, 2003 (24 months)	BMP-2 + allograft ring (N=14)	ICBG + allograft ring (N=12)	SF-36 PCS: MD 16.7 vs. MD 14.7, p>0.05 SF-36 MCS: MD 21.8 vs. MD 7.2, p>0.05
Kanna, 2021 (12 months)	Allograft + patient's blood + titanium cage (N=13)	Local graft + titanium cage (N=14)	2-item SF-12: MD 48.7 vs. MD 65.9, p<0.05

BMP-2 = bone morphogenetic protein; CI = confidence interval; ICBG = iliac crest bone graft; i-Factor = biologic bone graft made of a small peptide bound to an anorganic bone mineral; MCS = mental component score; MD = mean difference; NDI = Neck Disability Index; PCS = physical component score; PEEK = polyetheretherketone; SF = short form

3.10.3.3.4 Harms

There was low-strength evidence that the use of BMP-2 in cervical spine fusion is associated with increased complications compared to the use of no BMP-2 (SOE: Low), while evidence was inadequate to determine the comparative harms of other osteogenic materials (SOE: Insufficient).

Four RCTs (N=520) and 2 NRSI studies (N=944) reported harms with ACDF using various graft materials (**Table 8**). There were few differences between treatments reported in the randomized trials in the likelihood of various harms. One trial (N=319) reported a moderately greater likelihood of experiencing a new radiculopathy with an allograft ring with local graft than with i-Factor (13.66% vs. 25.00%, p=0.0142) but there were no differences in new intractable neck pain or progression of neuropathy.[133] One trial (N=100) reported a shorter hospital stay with a biphasic calcium phosphate ceramic combined with a PEEK cage compared with a PEEK cage with iliac crest bone graft.[135] Reasons for the difference in hospital stay were not provided.

Two retrospective NRSI of BMP-2 compared with no BMP-2 in ACDF (N=944) reported a greater likelihood of heterotopic ossification (78.6% vs. 59.2%, p<0.001)[138] and complications associated with neck swelling[139] with the use of BMP-2 (**Table 8**). In one NRSI, participants were 10 times more likely to have a neck swelling complication if BMP-2 was used in anterior cervical fusion, even after controlling for potential confounding variables (e.g., age, gender, presence of myelopathy, levels fused, smoking).[139]

3. Results

Table 8. Adverse events with ACDF using various graft materials

Trial (Timepoint)	Intervention A (Sample size)	Intervention B (Sample size)	Findings
Arnold, 2018 (24 months)	i-Factor + allograft ring (N=165)	Local graft + allograft ring (N=154)	Pseudarthrosis: 12.73% vs. 16.23%, p=0.3790 New intractable neck pain: 44.72% vs. 42.11%, p=0.1149 New radiculopathy: 13.66% vs. 25.00%, p=**0.0142** Adjacent segment degeneration: 13.04% vs. 16.45%, p=0.4274 Retropharyngeal hematoma/airway obstruction: 0% vs. 0.66%, p=0.4856 Progression of myelopathy: 0.62% vs. 0%, p=1.0 Additional cervical spine surgery: 7.45% vs. 10.53%, p=0.34
Baskin, 2003 (24 months)	BMP-2 + allograft ring (N=18)	ICBG + allograft ring (N=15)	Additional cervical spine surgery: 5.6% vs. 0%, p>0.05
Cho, 2005 (6 months)	Biphasic calcium phosphate ceramic + PEEK cage (N=50)	ICBG + PEEK cage (N=50)	Hospital stay (days): 4.43 vs. 7.00, **p=0.02**
Xie, 2015 (24 months)	Calcium sulphate + demineralized bone matrix + PEEK cage (N=35)	Autogenous iliac cancellous bone + PEEK cage (N=33)	Major complications: 0% vs. 0%, p=1.0 Additional cervical spine surgery: 0% vs. 0%, p=1.0
Arnold, 2016 (Retrospective; used propensity scoring)	BMP-2 + PEEK cage + titanium plate (N=224)	Cortical allograft ring + local bone + Atlantis Plate (N=486)	Heterotopic ossification 24 months postoperatively: 78.6% vs. 59.2%, **p<0.001**
Smucker, 2006 (Retrospective: adjusted for potential confounders)	BMP-2 (N=69)	No BMP-2 (N=165)	Neck swelling complications: 27.5% vs. 3.6%, **p<0.001** Delay in discharge: 13% vs. 3%, p=NR Severe dysphagia: 7% vs. 1%, p=NR Reintubation: 3% vs. 0%, p=NR PEG placement: 1% vs. 1%, p=NR Tracheostomy: 1% vs. 0.6%, p=NR Incision and drainage of swollen surgical site: 4% vs. 0%, p=NR Readmission to manage swelling: 3% vs. 0%, p=NR

BMP-2 = bone morphogenetic protein; ICBG = iliac crest bone graft; i-Factor = biologic bone graft made of a small peptide bound to an anorganic bone mineral; PEEK = polyetheretherketone; PEG = percutaneous endoscopic gastrostomy

3.11 Key Question 10. In patients with pseudarthrosis after prior anterior cervical fusion surgery, what are the comparative effectiveness and harms of posterior approaches compared to revision anterior arthrodesis?

No studies met eligibility criteria for Key Question 10.

3.12 Key Question 11. In patients with cervical spondylotic myelopathy, what is the prognostic utility of preoperative magnetic resonance imaging (MRI) findings for neurologic recovery after surgery?

3.12.1 Key Findings

- There was low-strength evidence that multisegmental T2-weighted-increased signal intensity (ISI) and sharp T2-weighted-increased signal intensity on preoperative MRI was associated with poorer outcomes (SOE: Low).
- There was low-strength evidence that increased signal intensity ratio (SIR) was associated with poorer neurologic recovery (SOE: Low).
- Evidence for other MRI findings was inadequate (SOE: Insufficient).

3.12.2 Description of Included Studies

MRI of the cervical spine is a common imaging procedure performed prior to cervical spine surgery. To identify whether MRI findings can predict neurologic recovery after surgery, we identified one relevant systematic review[140] (that included 22 studies) and 15 additional studies[141-156] that were not included in the systematic review or published subsequent to the review's search dates that provided evidence for this question (**Appendix C**). Studies were conducted in USA, China, Taiwan, United Kingdom, Spain, Italy, Greece, India, Korea, and Japan. Most studies were small, with sample sizes ranging from 19 to 861 (mean 108) participants. Mean age of participants ranged from 47 to 70 years (overall mean: 55.8 years), and the proportion of females ranged from 7 to 50 percent (mean 30%). The systematic review and 12 of the 15 primary studies were rated moderate risk of bias, with 3 studies rated as high risk of bias (**Appendix D**). Evidence was insufficient for MRI findings other than ISI and SIR due to limited available data for other outcomes (**Appendix G**).

3.12.3 Detailed Analysis

3.12.3.1 Fusion

No studies reported fusion outcomes.

3.12.3.2 Pain

No studies reported pain outcomes.

3.12.3.3 Function

3.12.3.3.1 Systematic Review Evidence

A 2013 systematic review that assessed the prognostic utility of preoperative magnetic resonance imaging (MRI) for neurologic recovery after surgery included 22 studies (N=1,508).[140] The included studies evaluated preoperative MRI in patients undergoing cervical disc surgery using a posterior approach (k=7), ACDF (k=5), mixed approaches (k=9), or an unspecified procedures (k=1) over followup ranging from 1.5 to 60.6 months (mean 27.8; standard deviation 4.6 months). The majority of patients in the included studies were male (mean

proportion of females: 27.1%), and the mean age (from 20 studies reporting age) was 57.4 (standard deviation, 1.0) years. Heterogeneity of study designs, methods, and outcomes (JOA in 17 studies, Nurick grade in 5 studies, Neck Disability Index in one study, and Neurosurgical Cervical Spine Score in one study) of the included studies precluded pooling of study findings, and the mixed results were reported narratively. Presence of multisegmental T2-weighted imaging signal intensity (ISI) was associated with worse functional outcomes in 5 studies, not associated with outcomes in 4 other studies, and lack of T2-weighted ISI was associated with better outcomes in 3 studies; qualitative classification of T2-weighted ISI was associated with poorer functional status in 6 studies, not associated with functional outcomes in 1 study, and lack of T2-weighted ISI associated with better outcomes in 1 study. Snake-eye appearance on axial T2-weighted MRI, ISI in gray and white matter, and increased signal intensity ratio (SIR) were associated with poorer surgical outcomes in one study each.

3.12.3.3.2 Primary Study Evidence

We identified four relevant studies (N=326) that were not included in the systematic review,[152,153,155,156] as well as 11 studies (in 12 publications) that were published subsequent to the review search dates.[141-151,154] Of these studies, two assessed presence of segmental abnormalities (endplate abnormalities, modic changes, and Cobb angle/loss of lordosis),[142,143,148] six assessed qualitative differences in ISI intensity,[141,145,146,150,153] three assessed SIR,[144,149,151] one evaluated presence or absence of signal changes,[155] one evaluated diffusion tensor tractography grading,[154] one evaluated the size of the transverse area at the compression site,[156] and one evaluated size, extent, and qualitative intensity.[147] The study (N=55) that assessed the size of the transverse area reported significant associations with postoperative JOA scores (r=0.298) and with JOA recovery (r=0.295) (both $p<0.05$).[156] The study (N=56) that evaluated size, extent, and intensity of ISI reported no association of size or extent of ISI with functional outcomes;[147] one other study of qualitative imaging signal intensity also reported no association of intensity changes with recovery (mJOA score ≥16, RR 1.71; 95% CI 0.90 to 3.24),[141] while four studies (N=714) did find qualitative intensity associated with reduced recovery ratio, lower likelihood of optimal surgical outcome, or change in JOA or NDI scores.[145-147,150] One study (N=52) reported improved JOA recovery rate (54.3% vs. 27.3%) in patients without ISI compared to those with ISI.[152] Another study (N=146) that assessed presence or absence of imaging signal changes reported that patients without imaging signal changes were more likely to have improvement in Nurick grade (OR 5.1; 95% CI, 1.87 to 25.1); however, there was no difference between patients without imaging signal changes and those with only T2-weighted signal changes.[155] Another study (N=73) found that the combination of T1-weighted hypointensity and T2-weighted hyperintensity was associated with poorer JOA recovery than T2-weighted hyperintensity alone or no ISI changes (JOA recovery 48% vs. 19% vs. 60.7%; T1- and T2-weighted ISI changes vs. T2-weighted ISI change only, p=0.0259).[153] Two studies of SIR (N=220) reported increased T2-weighted SIR associated with JOA recovery;[144,151] one study (N=148)[149] reported no association between T2-weighted SIR and outcomes, while lower T1-weighted SIR was associated with poorer neurological outcomes assessed with the JOA. One study (N=129)[154] found that diffusion tensor tractography grading using MRI images was associated with JOA score changes (r= -0.813, $p<0.001$) and JOA recovery (r= -0.429, $p<0.001$), while conventional MRI ISI grading was associated with JOA score changes (r= -0.674, $p<0.001$) but not with JOA recovery (r= -0.197, p=0.058).

3. Results

One study (N=121) reported a novel classification system for reporting loss of cervical lordosis following laminoplasty was predicted by an interplay of preoperative Cobb angle, T1 slope, and dynamic extension reserve.[148] One study (N=861) reported Modic changes, defined as "subchondral vertebral bone marrow lesions of the endplate" on preoperative MRI and found that while modic changes were associated with greater postoperative disability, modic changes were also associated with older age, greater number of levels fused, and a longer duration of symptoms.[142]

Comparing findings across studies was difficult due to the various study methods used (e.g., different type and basis of classification of T2W ISI [e.g., single segment, multisegment, L2 classification, Q3 classification, signal intensity ratio], different outcomes assessed [e.g., JOA, NDI, Nurick grade] and different methods to analyze the data [e.g., correlation, linear regression, multivariable regression, Student's *t* test]). Preoperative MRI also preceded different types of surgery (e.g., ACDF, laminoplasty, posterior-anterior decompression), which reduces the generalizability of findings.

3.12.3.3.3 Synthesis of Systematic Review and Primary Study Findings

There was low-strength evidence that multisegmental T2-weighted-increase signal intensity and sharp T2-2eighted-increased signal intensity on preoperative MRI was associated with poorer neurologic outcomes (SOE: Low); there was also low-strength evidence that increased signal intensity ratio of preoperative MRI was associated with poorer neurologic recovery (SOE: Low)

In total, presence of ISI was associated with poorer neurologic outcomes (e.g., JOA recovery, Nurick grade, NDI) in 7 studies and absence of ISI was associated with better neurologic outcomes (JOA, Nurick grade) in 4 studies but was not associated with changes in neurologic outcomes in 5 studies. Qualitative grading (increased intensity) of ISI was associated with worse neurologic outcomes (JOA, NDI) in 11 studies, absence of T2-weighted intensity associated with a better neurologic outcome (Nurick grade) in 1 study, and not associated with neurologic outcomes in 3 studies. Higher SIR was associated with poorer recovery in 3 studies (AUCs ranged from 78.6% to 87.3% in the two studies that reported accuracy results); one study reported lower SIR on T1-weighted associated with poorer neurological outcomes (JOA), while T2-weighted SIR was not associated with outcomes. One study reported that diffusion tensor tractography grading was more closely associated with neurological outcomes and recovery (JOA) than conventional ISI grading.

3.12.3.4 Quality of Life

No studies reported quality of life outcomes.

3.12.3.5 Harms

No studies reported harms or adverse events.

3.13 Key Question 12. What is the sensitivity and specificity of imaging assessment for identifying symptomatic pseudarthrosis after prior cervical fusion surgery?

3.13.1 Key Findings

- There is low-strength evidence that postoperative ACDF dynamic radiographs can predict pseudarthrosis in a largely asymptomatic population (SOE: Low) and a largely symptomatic population (SOE: Low).

3.13.2 Description of Included Studies

Two nonrandomized studies (N=722)[157,158] assessed diagnostic accuracy of radiographs in predicting pseudarthrosis after prior cervical fusion surgery (**Appendix C**). Both studies were conducted in the U.S. The mean ages of participants were 51 years and 54 years; the proportion of females were 54% and 62%. Neither study reported race or ethnicity. In both studies, enrolled patients had undergone ACDF as the index surgery, and revision surgery included anterior or posterior approaches.

Both studies were rated moderate risk of bias (**Appendix D**). Methodological limitations included lack of clarity on the number and characteristics of patients missing imaging studies. Neither study received funding.

3.13.3 Detailed Analysis

There is low-strength evidence that postoperative ACDF dynamic radiographs can predict pseudarthrosis in a largely asymptomatic and a largely symptomatic population (SOE: Low).

One study (N=125) reported diagnostic accuracy of dynamic radiographs and CT scans for identifying pseudarthrosis in patients who had undergone revision surgery for pseudarthrosis or adjacent segment pathology, using surgical exploration of fusion as the reference standard.[158] Medical records were retrospectively reviewed for patients operated on from January 2004 through December 2011. There were 262 levels evaluated (109 fused and 153 with pseudarthrosis). Most patients (84%) had revision surgery due to suspected pseudarthrosis, although it is unclear if patients were symptomatic. In dynamic radiographs magnified 150%, the optimal cutoff in interspinous motion to predict pseudarthrosis was 0.9 mm (AUC 0.899). Using cutoff criteria of interspinous motion ≥1 mm and superadjacent interspinous motion ≥4 mm resulted in similar values for diagnostic accuracy in dynamic radiographs versus a CT scan: sensitivity (86.3% vs. 87.2%), specificity (96.1% vs. 97.4%), positive predictive value (96.9% vs. 97.9%) and negative predictive value (83.4% vs. 84.4%).

One study (N=597, levels=1203) assessed diagnostic accuracy of dynamic radiographs for predicting symptomatic pseudarthrosis in patients who were largely asymptomatic but required revision surgery.[157] Medical records from 2010 to 2019 were reviewed for eligible patients. The reference standard was intraoperative documentation of pseudarthrosis (36% of the patient sample); only 4.9% of patients required pseudarthrosis revision.[157] Pseudarthrosis rates increased as the number of operative levels increased from 22.2% with 1-level to 75% with 4-level surgery. In radiographs taken one year post-primary surgery, using an optimal cutoff of 1 mm interspinous motion (AUC 0.868) had high negative predictive value (99.6%) and sensitivity (89.7%); moderate specificity (81%); and low positive predictive value (13.7%) in identifying

patients requiring revision surgery due to pseudarthrosis. Adding superadjacent interspinous motion ≥4 mm to 1 mm interspinous motion to the model, versus 1 mm alone,[158] reduced the number of patients and levels included in the authors' analysis but resulted in similar AUC. The positive predictive value was also decreased without improving the negative predictive value.

3.14 Key Question 13. In patients with cervical spondylotic myelopathy, what are the comparative effectiveness and harms of intraoperative neuromonitoring (e.g., with somatosensory or motor evoked potential measurements) versus no neuromonitoring on clinical outcomes in patients undergoing surgery?

3.14.1 Key Findings

- There was low-strength evidence of a similar likelihood of neurological complications with or without the use of intraoperative neuromonitoring (IONM) in ACDF (SOE: Low).

3.14.2 Description of Included Studies

Two retrospective NRSIs utilized large US claims databases (National Inpatient Sample [NIS]) of the Healthcare Cost and Utilization Project (HCUP) from 2009 to 2013 (N=141,007)[159] and PearlDiver from 2007 to 2014 (N=15,395)[160] to examine the effects of intraoperative neuromonitoring (IONM) versus no IONM in patients undergoing ACDF.

In the NIS study, 1:1 propensity score-matching, controlling for age, sex, indication, number of levels fused, Charlson Comorbidity Index (CCI) and admission type (elective, nonelective) was used (N=18,760).[159] There was no adjustment for confounders in the PearlDiver study.[160] The NIS data included inpatient data with no outpatient followup; the PearlDiver data included followup out to 30 days postoperatively. All data were collected from claims in the United States.

The mean age of participants was 54 years in the NIS study and reported by categories in the PearlDiver study (<45 years, 45-54, 55-64, 65-74, and >75; with the largest number of patients in the 45-54 age category). The average proportion of females was 51% and 52%, respectively. The NIS study enrolled a majority of White participants (80%), while the PearlDiver study did not report race/ethnicity (**Appendix C**).

Of patients with degenerative disease in the entire NIS, 42% of participants had radiculopathy alone and 31% had myelopathy (these proportions were not reported in the propensity score-matched NIS). Additionally, 66% of participants in the NIS study had a CCI of 0 (3.4% with a CCI of 3 or higher) and 84% had 1-2 level fusion, whereas the PearlDiver study did not report proportions with baseline radiculopathy, myelopathy, comorbidities, or levels fused.

The NIS study was rated moderate risk of bias due to study design.[159] The PearlDiver study was rated high risk of bias due to study design and lack of adjustment for potential confounders[160] (**Appendix D**).

3.14.3 Detailed Analysis

3.14.3.1 Outcomes

No studies reported fusion outcomes, pain, function, or quality of life.

3.14.3.2 Harms

There was low-strength evidence of a similar likelihood of neurological complications with or without the use of intraoperative neuromonitoring in ACDF (SOE: Low).

The NIS study included 18,760 patients who underwent ACDF in the propensity score-matched analyses from 2009 to 2013 and found no differences between IONM and no IONM in the rate of neurological complications (0.22% vs. 0.17%, p=0.41) or in the proportion of patients who required a hospital stay greater than 2 days (17.8% vs. 18.6%, p=0.15).[159]

The PearlDiver database study included 15,395 patients who underwent ACDF from 2007 to 2014 for degenerative radiculopathy or myelopathy (ION was used for 17.1% of patients, N=2627).[160] Although there was no propensity score matching or adjustments made for confounding variables, the results were similar to the NIS study. There was no difference in rate of neurologic complication within 30 days of the index procedure between IONM and no IONM (0.23% vs. 0.27%, p=0.84). However, younger patients were more likely to receive IONM (20.3% in patients less than 45 years of age compared to 13.6% in patients >75 years).

3.15 Contextual Question 1. What is the prevalence of cervical degenerative disease with spinal cord compression in asymptomatic patients?

Not all individuals with cervical degenerative disease that includes spinal cord compression (SCC) experience pain, radiculopathy, myelopathy or other symptoms. A 2021 systematic review and meta-analysis rated moderate risk of bias included 11 studies (N=3,686) that reported cervical MRI results in healthy individuals.[161] In pooled analysis, the prevalence of asymptomatic spinal cord compression was 24.2% (range 5.3% to 59%; 95% CI 12.4% to 36%, I^2=88).

To help explain the high statistical heterogeneity in pooled analysis, studies of asymptomatic participants were stratified based on mean age (less than or equal to 60 years versus greater than 60 years). The prevalence of SCC was lower in the younger subgroup (7 studies, N=1841, prevalence 7.4%, 95% CI 2.8% to 12%, I^2=40%) versus the older subgroup (4 studies, N=1845, prevalence 35.3%, 95% CI 14.1% to 56.5%, I^2=94%). Studies were also stratified based on study location: America/Europe (6 studies, N=390, prevalence of SCC 39.7%, 95% CI 21.0% to 58.3%, I^2=64%) versus Asia (5 studies, N=3296, prevalence of SCC 11.1%, 95% CI 1.6% to 20.5%, I^2=83%). The study with the largest number of participants (N=1211) was conducted in Japan, enrolled younger participants (mean age 50 years) and reported the lowest prevalence of SCC (5.3%).[162] In this study, spinal cord compression was defined as when "the AP diameter of the spinal canal at its narrowest was less than or equal to the AP diameter of the spinal cord at the C5 vertebral level."[162] This is in contrast to the study with the highest prevalence of participants with SCC (59%, N=183) that enrolled older participants (mean 66 years) and was conducted in the Czech Republic.[163] The definition of SCC in this study was more liberal and was diagnosed when "a change in spinal cord contour at the level of an intervertebral disc on axial or sagittal MRI compared with that at the midpoint level of neighboring vertebrae."[163] In both studies, as expected, the prevalence of SCC increased with age.

3.16 Contextual Question 2. What is the natural history of untreated spinal cord compression in patients with cervical degenerative disease?

The natural history of degeneration of the cervical spine progressing to non-myelopathic spinal cord compression (NMSCC) and ultimately cervical spondylotic myelopathy (CSM) is a continuum of disease that remains poorly understood. Untreated spinal cord compression is most studied in the context of CSM. There is a subset of patients with spinal cord compression found on imaging who are asymptomatic. A recent systematic review by Nouri et al (2022)[164] found the prevalence of asymptomatic spinal cord compression in healthy volunteers to be 24.2% (range 5.3 to 59%). A small series by Martin et al (2018)[165] looking at 20 asymptomatic patients with MRI evidence of spinal cord compression revealed that 2 (10%) developed symptoms of myelopathy at a median follow up of 21 months. The largest prospective study evaluating the transition from NMSCC to CSM by Bednarik et al (2008) revealed that among 199 patients enrolled with NMSCC, 8% developed CSM at 1 year follow up and 22.6% of patients developed CSM at median follow up of 44 months (range 1-12 years).[166] Factors found to independently predict the development of myelopathy in a multivariate analysis included presence of radiculopathy, spinal cord cross-sectional area and compression ratio.[167]

3. Results

CSM is the leading cause of spinal cord dysfunction among adults worldwide.[168] The pathogenesis of CSM is due to both mechanical and neuropathic changes to the spinal cord and blood spinal cord barrier (BSCB) generated by compression on the spinal cord.[169-172] The compressed cervical spinal cord is subjected to chronic hypoxic conditions due to dysfunction of endothelial cells as well as flattening and consequent loss of surrounding vessels.[170]

While the natural history of CSM in patients varies greatly, it is generally thought of as a progressive disorder. This was confirmed in a recent systematic review[173] that found moderate evidence from small prospective and retrospective studies that the proportion of patients who deteriorate by at least 1 point in the Japanese Orthopaedic Association (JOA) scale ranged from 20% to 60%. It is important to point out that these studies did not consider the minimal detectible difference to define deterioration, which is > 1 point based on reliability studies.[174,175] The overall lack of large, well designed and controlled studies evaluating the natural history of untreated spinal cord compression in patients with cervical degenerative diseases impairs clinicians' ability to counsel patients. A recent clinical practice guideline provided by the AO spine group suggested that either surgery or clinical observation are reasonable initial treatment options in mild CSM (e.g., mJOA score greater than or equal to 15).[176,177]

Shimomura et al[178] evaluated prognostic factors for deterioration of patients with CSM treated nonoperatively. Their prospective study included 56 patients with mild CSM, 11 (20%) had clinical deterioration over a mean follow up period of 35.6 months. Age, gender, follow up period, developmental or dynamic canal factors (e.g., canal size of < 12mm) of cervical spine on plane lateral radiographs, presence of high intensity of the cord on T2 weighted MRI and circumferential spinal cord compression on axial MRI were all evaluated as possible predictors for progression of myelopathy. However, they found the only predictive factor was presence of circumferential spinal cord compression on axial MRI (adjusted OR 26.6, 95% CI 1.7 to 421.5).[178] More studies are needed to better define the natural history of untreated spinal cord compression in the setting of degenerative changes along with predictors of progression.

4. Discussion

4.1 Findings in Relation to the Decisional Dilemmas

Cervical degenerative disease, which affects millions of older Americans, may lead to neck pain, radiculopathy, and myelopathy. Treatment of CDD, initially limited to conservative therapies (e.g., neck collar, traction, physiotherapy), has evolved to include instrumented and noninstrumented surgeries to decompress nerve roots and/or the spinal cord. Decisional dilemmas concerning best management of CDD include determination of whether one or more nonoperative treatments instead of surgery or in addition to surgery is preferred, and, if surgery is indicated, the determination of the most effective operative approaches and techniques for each individual patient. The key findings and strength of the evidence (SOE) are summarized in **Table 9**.

Fifty-six randomized trials (in 80 publications) and 49 nonrandomized studies (in 50 publications) and one systematic review provided evidence for this review. The highest quality evidence was for cervical disc arthroplasty versus ACDF in patients with cervical radiculopathy and/or myelopathy. Evidence for nonsurgical interventions was particularly limited. Similarly, there was no evidence to guide treatment for asymptomatic patients with radiographic spinal cord compression.

Conservative (nonoperative) therapy or operative treatment. There was insufficient evidence to determine the effectiveness of nonoperative compared with operative treatment for CDD, and limited evidence to suggest no important difference in pain beyond two weeks when a postoperative cervical collar was added to laminoplasty (SOE: Low). Post-operative pulsed electro-magnetic field stimulation in addition to ACDF was associated with a greater likelihood of fusion than ACDF alone (SOE: Low). Evidence for exercise therapy was insufficient.

Anterior or posterior surgery. Anterior approaches included anterior cervical foraminotomy, ACDF, and anterior decompression without fusion; posterior approaches included posterior cervical discectomy and fusion, laminoplasty and posterior cervical foraminotomy. Single-level surgery was performed in patients with radiculopathy and two or more levels in patients with myelopathy. There was no important difference between an anterior versus a posterior approach in pain and function in patients with CDD (SOE: Low). There was limited evidence to suggest that a posterior approach is associated with increased likelihood of experiencing any serious adverse event in patients with greater than or equal to 3-level disease (SOE: Low).

Laminoplasty or laminectomy and fusion. In patients with cervical spondylotic myelopathy, there was moderate strength evidence indicating similar benefits on postoperative function between laminectomy and fusion compared with laminoplasty and no important difference in reoperation rates, although limited evidence suggests laminoplasty may be associated with fewer complications than laminectomy and fusion (SOE: Low).

Disc replacement or fusion. In patients with radiculopathy and/or myelopathy at one level, there was moderate strength evidence of no important difference between C-ADR and ACDF in pain or function. C-ADR was associated with substantially decreased likelihood of reoperation (SOE: High) and slightly lower likelihood of any serious adverse event in the short term (SOE: Low), but there was no important difference between C-ADR and ACDF in serious adverse events longer term (SOE: Low). Study findings were similar in patients with 2-level C-ADR or ACDF in pain and function and likelihood of reoperation at the index level, but the likelihood of an adverse event was slightly lower at 24 with months with C-ADR and no different at 120

months (SOE: Low). Evidence was sparse for this comparison beyond two levels. The majority of these trials were industry funded.

In patients with pseudarthrosis after ACDF, evidence on comparative effectiveness and harms of revision anterior arthrodesis versus a posterior approach was lacking.

ACDF graft choices. In patients undergoing ACDF, there was moderate strength evidence of no important difference between use of a standalone cage or a plate and cage in fusion rate, postoperative arm pain, function, quality of life, or subsidence. In a comparison of titanium/titanium-coated cages versus PEEK cages in ACDF, there was limited evidence to suggest that use of a PEEK cage results in a greater likelihood of fusion and function improvement than use of a titanium/titanium-coated cage (SOE: Low). In patients undergoing ACDF, there was also low strength evidence to suggest an increased risk of complications with the use of BMP-2 in the cervical spine compared with fusion without the use of BMP-2 (i.e., use of other osteogenic materials).

Other decisional dilemmas included the use of pre- and post-operative imaging findings and associations with better or worse outcomes, and the use or nonuse of intraoperative neuromonitoring on patients undergoing cervical spine surgery.

Role of imaging. Evidence for imaging to predict neurologic recovery was heterogeneous, as various study methods were used (e.g., different type and basis of classification of increased signal intensity, different outcomes, and different statistical analysis methods), thus making comparisons across studies challenging. In patients with cervical myelopathy, there was limited evidence to suggest that multisegmental T2-weighted increased signal intensity, sharp T2-weighted increased signal intensity, and increased signal intensity ratio are associated with poorer neurologic recovery (SOE: Low).

In an asymptomatic and symptomatic populations, there was limited evidence suggesting that postoperative ACDF dynamic radiographs can predict pseudarthrosis with surgical exploration used as the gold standard (SOE: Low).

Intraoperative neuromonitoring or no monitoring. There was limited evidence to suggest that patients undergoing cervical spine surgery with IONM had similar likelihood of neurological complications as patients undergoing surgery without IONM (SOE: Low). Two databases (National Inpatient Sample [NIS] and PearlDiver) were included, but only the NIS analysis used propensity score matching. The PearlDiver study did not match or control for confounders, but had similar results. In the total NIS sample, 42 percent of participants had radiculopathy alone and 31 percent had myelopathy (proportions not reported in the matched sample), 66 percent had a Charlson Comorbidity Index of 0, and 84 percent had 1-2 level fusion. The PearlDiver study did not report baseline radiculopathy, myelopathy, comorbidities or levels fused.

Table 9. Summary of Findings: Cervical Degenerative Disease Treatment

Key Question	Comparison	Fusion *Effect Direction (SOE)*	Pain *Effect Direction (SOE)*	Function *Effect Direction (SOE)*	Quality of Life *Effect Direction (SOE)*	Adverse Events *Effect Direction (SOE)*
KQ 1. Radiographic and spinal cord compression and no myelopathy	Surgery vs. nonoperative treatment	No evidence	No evidence	No evidence	No evidence	No evidence

4. Discussion

Key Question	Comparison	Fusion *Effect Direction (SOE)*	Pain *Effect Direction (SOE)*	Function *Effect Direction (SOE)*	Quality of Life *Effect Direction (SOE)*	Adverse Events *Effect Direction (SOE)*
KQ 2. Radiographic spinal cord compression and mild to severe myelopathy	Surgery vs. nonoperative treatment	No evidence	No evidence	Insufficient evidence	No evidence	Insufficient evidence
KQ 3. In cervical degenerative disease	Surgery vs. nonoperative treatment	No evidence	Insufficient evidence	Insufficient evidence	No evidence	No evidence
KQ 4. In cervical degenerative disease	ACDF vs. ACDF + collar	Insufficient evidence	Insufficient evidence	Insufficient evidence	No evidence	No evidence
	ACDF vs. ACDF + electromagnetic stimulation (EMS)	Improved fusion rates favors EMS (SOE: Low)	Insufficient evidence	Insufficient evidence	No evidence	No evidence
	Laminoplasty vs. Laminoplasty + collar	Not applicable	No important difference (SOE: Low)	No important difference (SOE: Low)	No evidence	No evidence
	Laminoplasty vs. laminoplasty + exercise	Not applicable	Insufficient evidence	No evidence	No evidence	No evidence
KQ 5. In cervical radiculopathy	Anterior vs. posterior surgery	Insufficient evidence	<u>Neck and Arm pain</u>: No important difference (SOE: Low)	Insufficient evidence	Insufficient evidence	<u>Reoperation:</u> No important difference (SOE: Low)
KQ 6. In cervical degenerative disease with ≥3 level disease	Anterior vs. posterior surgery	Insufficient evidence	<u>Neck pain:</u> No important difference (SOE: Low) <u>Arm pain:</u> Insufficient evidence	No important difference (SOE: Low)	Insufficient evidence	<u>Mortality, severe dysphagia:</u> No important difference (SOE: Low) Reoperation (SOE: Insufficient) <u>Serious AE:</u> Moderate to Large favors anterior (SOE: Low)
KQ7. In cervical myelopathy	Laminectomy vs. Laminoplasty and fusion	No evidence	Insufficient evidence	No important difference (SOE: Moderate)	No evidence	<u>Reoperation:</u> No important difference (SOE: Moderate) <u>Adverse events:</u> Moderate to Large favors laminoplasty (SOE: Low)

4. Discussion

Key Question	Comparison	Fusion *Effect Direction (SOE)*	Pain *Effect Direction (SOE)*	Function *Effect Direction (SOE)*	Quality of Life *Effect Direction (SOE)*	Adverse Events *Effect Direction (SOE)*
KQ8. In cervical radiculopathy and/or myelopathy	Arthroplasty vs. ACDF	Not applicable	No important difference (SOE: Moderate)	No important difference (SOE: Moderate)	No evidence	Reoperation: High favors arthroplasty (1-level SOE: High) (2-level SOE: Low) Serious AE: Small favors arthroplasty (SOE: Low) Neurological events: No important difference (1-level SOE: Low) (2-level SOE: Insufficient)
KQ9. In ACDF	Standalone cage vs. plate and cage	No important difference (SOE: Moderate)	Neck pain: No important difference (SOE: Low) Arm pain: Insufficient evidence	No important difference (SOE: Low)	No important difference (SOE: Low)	Adjacent level ossification: No important difference (SOE: Low)
	Titanium/titanium-coated vs. PEEK cage	Small favoring PEEK (SOE: Low)	Insufficient evidence	Small favoring PEEK (SOE: Low)	No evidence	Insufficient evidence
	Autograft vs. allograft vs. other osteogenic materials	Insufficient Evidence	Insufficient evidence	Insufficient evidence	Insufficient evidence	Adverse events: Large favors nonuse of BMP-2 (SOE: Low)
KQ 10. In pseudarthrosis after prior anterior fusion surgery	Posterior approach vs. revision anterior arthrodesis	No evidence	No evidence	No evidence	No evidence	No evidence
KQ 11. In cervical myelopathy, prognostic utility of MRI for neurologic recovery	T2-weighted increased signal intensity and intensity ratio, sharp signal intensity	No evidence	No evidence	No evidence	No evidence	Neurologic recovery: favors no signal, less sharp signal, increased signal intensity ratio (SOE: Low)
KQ 12. Imaging to detect pseudarthrosis	Dynamic radiographs (asymptomatic population)	Predicts pseudarthrosis (SOE: Low)	Not applicable	Not applicable	Not applicable	No applicable
	Dynamic radiographs (symptomatic population)	Predicts pseudarthrosis (SOE: Low)	Not applicable	Not applicable	Not applicable	Not applicable

Key Question	Comparison	Fusion *Effect Direction (SOE)*	Pain *Effect Direction (SOE)*	Function *Effect Direction (SOE)*	Quality of Life *Effect Direction (SOE)*	Adverse Events *Effect Direction (SOE)*
KQ 13. In cervical myelopathy	IONM vs. no IONM in ACDF	No evidence	No evidence	No evidence	No evidence	Neurologic complications: No important difference (SOE: Low)

Effect Direction: none, slight/small, moderate, or large effect/improvement
Strength of Evidence: low, moderate, high
none = no effect/no statistically significant effect

4.2 Implications for Clinical and Policy Decisions

This review was sponsored by the Congress of Neurological Surgeons to update their 2009 guidelines on the management of cervical degenerative disease. Our review provides additional evidence that operative approaches to management of cervical degenerative disease generally result in some improvement in pain, function, and quality of life postoperatively, as well as successful fusion (if a fusion surgery). In many cases differences in patient-centered benefit outcomes between compared operative approaches and techniques were minimal. The likelihood of general or specific adverse events, such as need for reoperation/revision surgery, were where most differences between therapies were observed and may help guide decision making regarding best operative approach for any given patient.

Our review provides additional support to the 2009 finding that preoperative MRI can help predict better or worse outcomes and to the 2009 recommendation discouraging use of BMP-2 in the cervical spine. Standalone cages for cervical fusion represent a newer design (Zero-P approved for use in the U.S. in 2008) and not covered in the 2009 guidelines. Although a more modern design, we did not find it superior to the use of anterior plating for most outcomes.

Gaps in the evidence make it difficult to create recommendations and inform policy. For example, challenges remain in determining the preferred course of action in patients with incidental findings of spinal cord compression on MRI. Although the natural history of non-myelopathic spinal cord compression is poorly understood, limited evidence suggests that some patients develop myelopathy over time, but it is not clear if any treatment provided prior to the development of symptoms results in better outcomes than treating symptomatic disease. Another challenge remaining is determining when conservative treatment may be preferred and what therapies are most effective compared with operative management or result in better outcomes when added to surgery. Good quality comparative evidence on conservative treatment was sparse in this review.

4.3 Strength and Limitations of the Systematic Review Process

Strengths. This review appears to provide the most comprehensive synthesis of evidence related to the comparative effectiveness of surgical treatment of CDD and identifies important gaps in the comparative evidence for many of them. Important strengths of this review include the use of a "best evidence" approach, where we focused our efforts on studies with least risk of bias, particularly randomized trials when available and supplemented with nonrandomized studies that adjusted for potential prognostic variables where appropriate. We avoided use of nonrandomized studies that did not provide some means of adjustment (e.g., propensity score

matching, statistical control for confounding variables) as the conclusion from such studies may differ from RCT evidence and are more likely to suffer from various important biases (see below). Another strength is our focus on outcomes of primary importance to patients including pain, function, and quality of life as improved patient outcomes may lead to higher quality patient care, as well as patient satisfaction with care. Additionally, interpretation of clinically important differences in mean change for continuous variables is challenging. A strength of our review is our categorization of the magnitude of effects for function and pain outcomes using the system described in our previous reviews to facilitate interpretation of results across trials and interventions by providing a level of consistency and objective benchmarks for comparison. We also added two contextual questions (on the natural history of untreated spinal cord compression and on the prevalence of CDD with spinal cord compression in asymptomatic patients) to inform this review.

Limitations. For many KQs, quantitative synthesis of evidence was not possible due to the poor quality of evidence available and lack of comparative evidence for some key questions. For some key questions evidence was limited to one study per comparison, making it difficult to draw conclusions about any specific treatment. While we did include NRSIs that made comparisons of interest, results from such studies should be interpreted cautiously. Limitations of these studies generally led to determination of insufficient evidence for many outcomes. Confounding by indication, lack of adequate control for confounding on important prognostic factors, as well as failure to adequately account for selection of patients and loss to follow-up in NRSIs were common methodologic concerns. For subjective patient-reported outcomes such as pain, NRSI results may be misleading due to the subjective nature of pain and the impact of nonspecific effects related to patient expectations regarding treatment and attention received. Analysis of data from large administrative claims-based databases present additional methodological challenges. Coding related to conditions, procedures and outcomes in such databases is focused on optimizing billing and there is a potential for misclassification of exposures and outcomes. Such databases are unable to account for some potential confounders or for factors that may impact decision-making regarding the appropriateness of a given procedure, e.g., use of an anterior versus posterior procedure. The large sample sizes available for administrative data may facilitate evaluation of rare outcomes and may demonstrate statistical significance when results may be of unclear clinical importance.

Other limitations of our review include the following:

1) lack of RCT data for many comparisons and small sample sizes in most trials that precluded analyses on differential effectiveness and harms of interventions based on patient demographics, social determinants of health, severity of radiculopathy or myelopathy, number of vertebral levels involved, and other factors;

2) poor reporting of adverse events in many studies and heterogeneity in what harms and adverse events were described;

3) studies reporting vertebral levels affected (e.g., number of levels with pseudarthrosis, subsidence, needing reoperation) while not reporting the number of individuals experiencing a specific adverse event such as pseudarthrosis, thereby limiting the ability to use such studies in a pooled analysis in conjunction with studies reporting results in people rather than vertebral levels;

4) heterogeneity in research design, interventions, and reported outcomes for several key questions that limit ability to draw conclusions on effectiveness across studies;

5) in most cases we were not able to assess for publication bias using graphical or statistical methods to evaluate any potential impact of small sample sizes due to insufficient number of studies per comparison; and

6) limiting the evidence to English-language publications is a potential limitation, however we did not identify large numbers of non-English-language articles in our review of bibliographies.

4.4 Applicability

According to a NIS trend study of patients who underwent cervical fusion in 2013 for cervical spondylotic myelopathy (N=8181), the average patient was 60.6 years, slightly more likely to be male (54.3%), White (71.5%), with a CCI ≤ 2 (65.7%), have Medicare (44.6%) or private insurance (39.6%), and live in the South (43.8%).[179] In the absence of more recent data, this represents a "best guess" at defining the typical patient seen in clinical practice today. There were similarities and differences between the typical study participant in our review and the typical patient as described above.

Reasons for greater applicability of this body of evidence to clinical practice include: 1) many studies required enrolled study participants to have failed several weeks or months of conservative therapies, which is considered a valid approach to the management of mild degenerative cervical myelopathy (as is an operative approach);[176] 2) studies enrolled a balance of males and females; 3) most studies did not limit the upper age of enrollment and included individuals in their 60s or 70s (although the mean age of participants in most studies was in the 40s and 50s); and 4) studies often enrolled patients with a combination of radiculopathy and myelopathy, likely reflecting the condition of many U.S. patients. Additionally, approximately 45 percent of studies included in this review were conducted in the U.S.

Reasons for lower applicability to clinical practice include the exclusion of participants with a variety of common health conditions such as inflammatory arthritis, obesity, and diabetes. The risk of CDD increases with age and so do many other health conditions and comorbidities. For example, a large proportion of the U.S. population is overweight or obese and an increasing proportion have diabetes. Excluding these populations from surgical intervention studies, because postoperative improvement may be reduced, decreases the applicability of study findings to many U.S. patients needing operative management of their CDD. Additionally, few studies reported race or ethnicity. While those that did tended to enroll white participants, it is unclear how differences in access in populations of color may impact results.

4.5 Future Research

While it may not always be feasible to perform RCTs for surgical procedures, well-designed prospective comparative NRSIs with protocols using methods for patient selection and treatment allocation that mitigate possible selection bias and imbalances in prognostic factors and that follow protocols established *a priori* for comparable evaluation, measurement and treatment of groups would provide a valuable contribution to the evidence base. In order to evaluate the differential impact of patient characteristics and other factors, adequately powered RCTs are needed. Additionally, more explicit evaluation of procedure-specific (or device-specific) harms and adverse events is needed in future studies; ideally such studies would be powered to detect rare events. Future studies should also report the proportion of patients who experience a clinically important improvement in pain or function. This would provide valuable insight to complement data on average changes in continuous measures of pain, function, and quality of

life for which there is difficulty describing clinically important effects. Studies should also estimate the minimally important between-group differences for included outcomes to facilitate interpretation of study findings.

4.6 Conclusions

There were few differences in benefits between surgical approaches and techniques compared in included studies for the treatment of cervical degenerative disease. However, there were some differences in the frequency of adverse events for some comparisons. There was substantial evidence that the risk of reoperation is much lower for artificial disc replacement than ACDF. Limited evidence also suggests a lower likelihood of experiencing any serious adverse event with ACDF than PCDF and a lower risk for any complication with laminoplasty compared with laminectomy and fusion. There was limited evidence on the role of nonoperative management instead of surgery or in addition to surgery to treat CDD, and no evidence to determine benefits and harms of a revision anterior arthrodesis or posterior approach in patients with pseudarthrosis after prior anterior cervical fusion.

References

1. Matz PG, Anderson PA, Kaiser MG, et al. Introduction and methodology: guidelines for the surgical management of cervical degenerative disease. J Neurosurg Spine. 2009 Aug;11(2):101-3. doi: 10.3171/2009.1.SPINE08712. PMID: 19769488.
2. Fehlings MG, Arvin B. Surgical management of cervical degenerative disease: the evidence related to indications, impact, and outcome. J Neurosurg Spine. 2009 Aug;11(2):97-100. doi: 10.3171/2009.5.SPINE09210. PMID: 19769487.
3. Congress of Neurological Surgeons. Guideline for the Surgical Management of Cervical Degenerative Disease. Congress of Neurological Surgeons; 2009. https://www.cns.org/guidelines/browse-guidelines-detail/surgical-management-of-cervical-degenerative-disea2022.
4. Agency for Healthcare Research and Quality. Methods guide for effectiveness and comparative effectiveness reviews Agency for Healthcare Research and Quality. Rockville, MD: 2020. https://effectivehealthcare.ahrq.gov/products/collections/cer-methods-guide.
5. Parenteau CS, Lau EC, Campbell IC, et al. Prevalence of spine degeneration diagnosis by type, age, gender, and obesity using Medicare data. Sci Rep. 2021 Mar 8;11(1):5389. doi: 10.1038/s41598-021-84724-6. PMID: 33686128.
6. Buser Z, Ortega B, D'Oro A, et al. Spine degenerative conditions and their treatments: national trends in the United States of America. Global spine j. 2018 Feb;8(1):57-67. doi: 10.1177/2192568217696688. PMID: 29456916.
7. Schiedo RM, Narain A, Adams S, et al. 101. Prospective evaluation of degenerative cervical myelopathy in asymptomatic patients over 60 years. Spine Journal. 2020;20(9):S50-S1. doi: 10.1016/j.spinee.2020.05.207.
8. Banerjee A, Mowforth OD, Nouri A, et al. The prevalence of degenerative cervical myelopathy-related pathologies on magnetic resonance imaging in healthy/asymptomatic individuals: a meta-analysis of published studies and comparison to a symptomatic cohort. J Clin Neurosci. 2022 2022/05/01/;99:53-61. doi: 10.1016/j.jocn.2022.03.002. PMID: 35255357.
9. Broekema AEH, Groen RJM, Simoes de Souza NF, et al. Surgical interventions for cervical radiculopathy without myelopathy: a systematic review and meta-analysis. J Bone Joint Surg Am. 2020 Dec 16;102(24):2182-96. doi: 10.2106/JBJS.20.00324. PMID: 32842045.
10. Mummaneni PV, Kaiser MG, Matz PG, et al. Cervical surgical techniques for the treatment of cervical spondylotic myelopathy. J Neurosurg Spine. 2009 Aug;11(2):130-41. doi: 10.3171/2009.3.SPINE08728. PMID: 19769492.
11. Carette S, Fehlings MG. Clinical practice. Cervical radiculopathy. N Engl J Med. 2005 Jul 28;353(4):392-

9. doi: 10.1056/NEJMcp043887. PMID: 16049211.

12. Yoshimatsu H, Nagata K, Goto H, et al. Conservative treatment for cervical spondylotic myelopathy. prediction of treatment effects by multivariate analysis. Spine J. 2001 Jul-Aug;1(4):269-73. doi: 10.1016/s1529-9430(01)00082-1. PMID: 14588331.

13. Nouri A, Martin AR, Mikulis D, et al. Magnetic resonance imaging assessment of degenerative cervical myelopathy: a review of structural changes and measurement techniques. Neurosurg Focus; 2016. p. E5.

14. Boden SD, McCowin PR, Davis DO, et al. Abnormal magnetic-resonance scans of the cervical spine in asymptomatic subjects. A prospective investigation. J Bone Joint Surg Am. 1990 Sep;72(8):1178-84. PMID: 2398088.

15. Furlan AD, Malmivaara A, Chou R, et al. 2015 Updated Method Guideline for Systematic Reviews in the Cochrane Back and Neck Group. Spine (Phila Pa 1976). 2015 Nov;40(21):1660-73. doi: 10.1097/BRS.0000000000001061. PMID: 26208232.

16. US Preventive Services Task Force. Methods and Processes. Rockville, MD: 2019. https://www.uspreventiveservicestaskforce.org/uspstf/about-uspstf/methods-and-processes Accessed December 10, 2021.

17. Hardy RJ, Thompson SG. A likelihood approach to meta-analysis with random effects. Stat Med. 1996 Mar 30;15(6):619-29. doi: 10.1002/(sici)1097-0258(19960330)15:6<619::Aid-sim188>3.0.Co;2-a. PMID: 8731004.

18. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ. 2003 Sep 6;327(7414):557-60. doi: 10.1136/bmj.327.7414.557. PMID: 12958120.

19. Sterne JA, Sutton AJ, Ioannidis JP, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ. 2011 Jul 22;343:d4002. doi: 10.1136/bmj.d4002. PMID: 21784880.

20. Kadanka Z, Mares M, Bednaník J, et al. Approaches to spondylotic cervical myelopathy: conservative versus surgical results in a 3-year follow-up study. Spine (Phila Pa 1976). 2002 Oct 15;27(20):2205-10; discussion 10-1. doi: 10.1097/01.Brs.0000029255.77224.Bb. PMID: 12394893.

21. Kadanka Z, Mares M, Bednarík J, et al. Predictive factors for mild forms of spondylotic cervical myelopathy treated conservatively or surgically. Eur J Neurol. 2005 Jan;12(1):16-24. doi: 10.1111/j.1468-1331.2004.00947.x. PMID: 15613142.

22. Kadanka Z, Bednařík J, Novotný O, et al. Cervical spondylotic myelopathy: conservative versus surgical treatment after 10 years. Eur Spine J. 2011 Sep;20(9):1533-8. doi: 10.1007/s00586-011-1811-9. PMID: 21519928.

23. Colamaria A, Ciappetta P, Fochi NP, et al. Anterior cervical corpectomy for treatment of spondylotic

myelopathy. Results of a prospective double-armed study with a three-year follow-up. J Neurosurg Sci. 2022 Apr 13;13:13. doi: 10.23736/S0390-5616.22.05608-9. PMID: 35416453.

24. Persson LC, Lilja A. Pain, coping, emotional state and physical function in patients with chronic radicular neck pain. A comparison between patients treated with surgery, physiotherapy or neck collar--a blinded, prospective randomized study. Disabil Rehabil. 2001 May 20;23(8):325-35. doi: 10.1080/09638280010005567. PMID: 11374522.

25. Persson LC, Moritz U, Brandt L, et al. Cervical radiculopathy: pain, muscle weakness and sensory loss in patients with cervical radiculopathy treated with surgery, physiotherapy or cervical collar. A prospective, controlled study. Eur Spine J. 1997;6(4):256-66. doi: 10.1007/bf01322448. PMID: 9294750.

26. Abbott A, Halvorsen M, Dedering A. Is there a need for cervical collar usage post anterior cervical decompression and fusion using interbody cages? A randomized controlled pilot trial. Physiother. 2013 May;29(4):290-300. doi: 10.3109/09593985.2012.731627. PMID: 23074995.

27. Cheung JPY, Cheung PWH, Law K, et al. Postoperative rigid cervical collar leads to less axial neck pain in the early stage after open-door laminoplasty-a single-blinded randomized controlled trial. Neurosurgery. 2019;85(3):325-34p. doi: 10.1093/neuros/nyy359.

28. Foley KT, Mroz TE, Arnold PM, et al. Randomized, prospective, and controlled clinical trial of pulsed electromagnetic field stimulation for cervical fusion. Spine J. 2008 May-Jun;8(3):436-42. doi: 10.1016/j.spinee.2007.06.006. PMID: 17983841.

29. Hida T, Sakai Y, Ito K, et al. Collar fixation is not mandatory after cervical laminoplasty: a randomized controlled trial. Spine. 2017 Mar;42(5):E253-E9. doi: 10.1097/BRS.0000000000001994. PMID: 27879567.

30. Uehara T, Tsushima E, Yamada S, et al. A randomized controlled trial for the intervention effect of early exercise therapy on axial pain after cervical laminoplasty. Spine surgery and related research. 2022;6(2):123-32. doi: 10.22603/SSRR.2021-0110.

31. Ebrahim KS, El-Shehaby A, Darwish A, et al. Anterior or posterior foraminotomy for unilateral cervical radiculopathy. Pan arab journal of neurosurgery. 2011;15(2):34-46p.

32. Ruetten S, Komp M, Merk H, et al. Full-endoscopic cervical posterior foraminotomy for the operation of lateral disc herniations using 5.9-mm endoscopes: a prospective, randomized, controlled study. Spine (Phila Pa 1976). 2008 Apr 20;33(9):940-8. doi: 10.1097/BRS.0b013e31816c8b67. PMID: 18427313.

33. Wirth FP, Dowd GC, Sanders HF, et al. Cervical discectomy. A prospective analysis of three operative techniques. Surg Neurol. 2000 Apr;53(4):340-6; discussion 6-8. doi: 10.1016/s0090-3019(00)00201-9. PMID: 10825519.

34. Alvin MD, Lubelski D, Abdullah KG, et al. Cost-utility analysis of anterior cervical discectomy and fusion with plating (ACDFP) versus posterior cervical foraminotomy (PCF) for patients with single-level cervical radiculopathy at 1-year follow-up. Clin Spine Surg. 2016 Mar;29(2):E67-72. doi: 10.1097/BSD.0000000000000099. PMID: 26889994.

35. Lubelski D, Healy AT, Silverstein MP, et al. Reoperation rates after anterior cervical discectomy and fusion versus posterior cervical foraminotomy: a propensity-matched analysis. Spine J. 2015 Jun 01;15(6):1277-83. doi: 10.1016/j.spinee.2015.02.026. PMID: 25720729.

36. Foster MT, Carleton-Bland NP, Lee MK, et al. Comparison of clinical outcomes in anterior cervical discectomy versus foraminotomy for brachialgia. Br J Neurosurg. 2019 Feb;33(1):3-7. doi: 10.1080/02688697.2018.1527013. PMID: 30450995.

37. Witiw CD, Smieliauskas F, O'Toole JE, et al. Comparison of anterior cervical discectomy and fusion to posterior cervical foraminotomy for cervical radiculopathy: utilization, costs, and adverse events 2003 to 2014. Neurosurgery. 2019 02 01;84(2):413-20. doi: 10.1093/neuros/nyy051. PMID: 29548034.

38. Jiang YQ, Li XL, Zhou XG, et al. A prospective randomized trial comparing anterior cervical discectomy and fusion versus plate-only open-door laminoplasty for the treatment of spinal stenosis in degenerative diseases. Eur Spine J. 2017 04;26(4):1162-72. doi: 10.1007/s00586-016-4878-5. PMID: 27885472.

39. Asher AL, Devin CJ, Kerezoudis P, et al. Comparison of outcomes following anterior vs posterior fusion surgery for patients with degenerative cervical myelopathy: an analysis from quality outcomes database. Neurosurgery. 2019 04 01;84(4):919-26. doi: 10.1093/neuros/nyy144. PMID: 29741718.

40. Badhiwala JH, Ellenbogen Y, Khan O, et al. Comparison of the inpatient complications and health care costs of anterior versus posterior cervical decompression and fusion in patients with multilevel degenerative cervical myelopathy: a retrospective propensity score-matched analysis. World Neurosurg. 2020 Feb;134:e112-e9. doi: 10.1016/j.wneu.2019.09.132. PMID: 31574327.

41. Cole T, Veeravagu A, Zhang M, et al. Anterior versus posterior approach for multilevel degenerative cervical disease: a retrospective propensity score-matched study of the marketscan database. Spine. 2015 Jul 01;40(13):1033-8. doi: 10.1097/BRS.0000000000000872. PMID: 25768690.

42. Fehlings MG, Barry S, Kopjar B, et al. Anterior versus posterior surgical approaches to treat cervical spondylotic myelopathy: outcomes of the prospective multicenter aospine north america csm study in 264 patients. Spine. 2013;38(26):2247-52. doi: 10.1097/BRS.0000000000000047. PMID: 24108289.

43. Joo PY, Jayaram RH, McLaughlin WM, et al. Four-level anterior versus posterior cervical fusions: perioperative outcomes and five-year reoperation rates: outcomes after four-level anterior versus posterior cervical procedures. N Am Spine Soc J. 2022 Jun;10:100115. doi: 10.1016/j.xnsj.2022.100115. PMID: 35392022.

44. Lee NJ, Boddapati V, Mathew J, et al. What is the impact of surgical approach in the treatment of degenerative cervical myelopathy in patients with OPLL? A propensity-score matched, multi-center analysis on inpatient and post-discharge 90-day outcomes. Global spine j. 2021 Feb 19:2192568221994797. doi: 10.1177/2192568221994797. PMID: 33601898.

45. Lee NJ, Kim JS, Park P, et al. A comparison of various surgical treatments for degenerative cervical myelopathy: a propensity score matched analysis. Global spine j. 2022 Jul;12(6):1109-18. doi: 10.1177/2192568220976092. PMID: 33375849.

46. Nunna RS, Khalid S, Chiu RG, et al. Anterior vs posterior approach in multilevel cervical spondylotic myelopathy: a nationwide propensity-matched analysis of complications, outcomes, and narcotic use. Int J Spine Surg. 2022 Feb;16(1):88-94. doi: 10.14444/8198. PMID: 35314510.

47. Wadhwa H, Sharma J, Varshneya K, et al. Anterior cervical discectomy and fusion versus laminoplasty for multilevel cervical spondylotic myelopathy: a national administrative database analysis. World Neurosurg. 2021 08;152:e738-e44. doi: 10.1016/j.wneu.2021.06.064. PMID: 34153482.

48. Elmallawany M, Kandel H, Soliman MAR, et al. The safety and efficacy of cervical laminectomy and fusion versus cervical laminoplasty surgery in degenerative cervical myelopathy: a prospective randomized trial. Open Access Maced J Med Sci. 2020;8:807-14. doi: 10.3889/oamjms.2020.4841.

49. Manzano GR, Casella G, Wang MY, et al. A prospective, randomized trial comparing expansile cervical laminoplasty and cervical laminectomy and fusion for multilevel cervical myelopathy. Neurosurgery. 2012 Feb;70(2):264-77. doi: 10.1227/NEU.0b013e3182305669. PMID: 22251974.

50. Blizzard DJ, Caputo AM, Sheets CZ, et al. Laminoplasty versus laminectomy with fusion for the treatment of spondylotic cervical myelopathy: short-term follow-up. Eur Spine J. 2017 01;26(1):85-93. doi: 10.1007/s00586-016-4746-3. PMID: 27554354.

51. Fehlings MG, Santaguida C, Tetreault L, et al. Laminectomy and fusion versus laminoplasty for the treatment of degenerative cervical myelopathy: results from the AOSpine North America and International prospective multicenter studies. Spine J. 2017 01;17(1):102-8. doi: 10.1016/j.spinee.2016.08.019. PMID: 27597512.

52. He X, Zhang JN, Liu TJ, et al. Is laminectomy and fusion the better choice than laminoplasty for multilevel cervical myelopathy with

signal changes on magnetic resonance imaging? A comparison of two posterior surgeries. BMC Musculoskelet Disord. 2020 Jul 02;21(1):423. doi: 10.1186/s12891-020-03435-7. PMID: 32615953.

53. McDonald CL, Hershman SH, Hogan W, et al. Cervical laminoplasty versus posterior laminectomy and fusion: trends in utilization and evaluation of complication and revision surgery rates. J Am Acad Orthop Surg. 2022 May 30;30(17):30. doi: 10.5435/JAAOS-D-22-00106. PMID: 35640093.

54. Mesregah MK, Formanek B, Liu JC, et al. Perioperative complications of surgery for degenerative cervical myelopathy: a comparison between 3 procedures. Global spine j. 2021 Mar 12:2192568221998306. doi: 10.1177/2192568221998306. PMID: 33709809.

55. Woods BI, Hohl J, Lee J, et al. Laminoplasty versus laminectomy and fusion for multilevel cervical spondylotic myelopathy. Clin Orthop. 2011 Mar;469(3):688-95. doi: 10.1007/s11999-010-1653-5. PMID: 21089002.

56. Anderson PA, Sasso RC, Riew KD. Comparison of adverse events between the Bryan artificial cervical disc and anterior cervical arthrodesis. Spine. 2008 May 20;33(12):1305-12. doi: 10.1097/BRS.0b013e31817329a1. PMID: 18496341.

57. Burkus JK, Haid RW, Traynelis VC, et al. Long-term clinical and radiographic outcomes of cervical disc replacement with the Prestige disc: results from a prospective randomized controlled clinical trial. J Neurosurg Spine. 2010 Sep;13(3):308-18. doi: 10.3171/2010.3.SPINE09513. PMID: 20809722.

58. Burkus JK, Traynelis VC, Haid RW, Jr., et al. Clinical and radiographic analysis of an artificial cervical disc: 7-year follow-up from the Prestige prospective randomized controlled clinical trial: Clinical article. J Neurosurg Spine. 2014 Oct;21(4):516-28. doi: 10.3171/2014.6.SPINE13996. PMID: 25036218.

59. Chen X, Shi L, Yu X, et al. Comparative study of artificial cervical disc replacement and anterior cervical discectomy/fusion in the treatment of cervical spondylotic myelopathy. Int J Clin Exp Med. 2019;12(8):10597-604p.

60. Cheng L, Nie L, Li M, et al. Superiority of the Bryan(R) disc prosthesis for cervical myelopathy: a randomized study with 3-year followup. Clin Orthop. 2011 Dec;469(12):3408-14. doi: 10.1007/s11999-011-2039-z. PMID: 21997779.

61. Cheng L, Nie L, Zhang L, et al. Fusion versus Bryan Cervical Disc in two-level cervical disc disease: a prospective, randomised study. Int Orthop. 2009 Oct;33(5):1347-51. doi: 10.1007/s00264-008-0655-3. PMID: 18956190.

62. Cincu R, Lorente Fde A, Gomez J, et al. Long term preservation of motion with artificial cervical disc implants: a comparison between cervical disc replacement and rigid fusion with cage. Asian J Neurosurg. 2014 Oct-

Dec;9(4):213-7. doi: 10.4103/1793-5482.146608. PMID: 25685218.

63. Davis RJ, Kim KD, Hisey MS, et al. Cervical total disc replacement with the Mobi-C cervical artificial disc compared with anterior discectomy and fusion for treatment of 2-level symptomatic degenerative disc disease: a prospective, randomized, controlled multicenter clinical trial: clinical article. J Neurosurg Spine. 2013 Nov;19(5):532-45. doi: 10.3171/2013.6.SPINE12527. PMID: 24010901.

64. Davis RJ, Nunley PD, Kim KD, et al. Two-level total disc replacement with Mobi-C cervical artificial disc versus anterior discectomy and fusion: a prospective, randomized, controlled multicenter clinical trial with 4-year follow-up results. J Neurosurg Spine. 2015 Jan;22(1):15-25. doi: 10.3171/2014.7.SPINE13953. PMID: 25380538.

65. Delamarter RB, Murrey D, Janssen ME, et al. Results at 24 months from the prospective, randomized, multicenter Investigational Device Exemption trial of ProDisc-C versus anterior cervical discectomy and fusion with 4-year follow-up and continued access patients. Sas J. 2010;4(4):122-8. doi: 10.1016/j.esas.2010.09.001. PMID: 25802660.

66. Delamarter RB, Zigler J. Five-year reoperation rates, cervical total disc replacement versus fusion, results of a prospective randomized clinical trial. Spine. 2013 Apr 20;38(9):711-7. doi: 10.1097/BRS.0b013e3182797592. PMID: 23124255.

67. Donk RD, Verbeek ALM, Verhagen WIM, et al. What's the best surgical treatment for patients with cervical radiculopathy due to single-level degenerative disease? A randomized controlled trial. PLoS ONE. 2017;12(8):e0183603. doi: 10.1371/journal.pone.0183603. PMID: 28850600.

68. Ghobrial GM, Lavelle WF, Florman JE, et al. Symptomatic adjacent level disease requiring surgery: Analysis of 10-year results from a prospective, randomized, clinical trial comparing cervical disc arthroplasty to anterior cervical fusion. Neurosurgery. 2018 02 01;84(2):347-54. doi: 10.1093/neuros/nyy118. PMID: 29635520.

69. Gornet MF, Lanman TH, Burkus JK, et al. Two-level cervical disc arthroplasty versus anterior cervical discectomy and fusion: 10-year outcomes of a prospective, randomized investigational device exemption clinical trial. J Neurosurg Spine. 2019 Jun 21;31(4):1-11. doi: 10.3171/2019.4.SPINE19157. PMID: 31226684.

70. Gornet MF, Lanman TH, Burkus JK, et al. Cervical disc arthroplasty with the Prestige LP disc versus anterior cervical discectomy and fusion, at 2 levels: results of a prospective, multicenter randomized controlled clinical trial at 24 months. J Neurosurg Spine. 2017 Jun;26(6):653-67. doi: 10.3171/2016.10.SPINE16264. PMID: 28304237.

71. Gornet M, McConnell J, Riew K, et al. Treatment of cervical myelopathy: long-term outcomes of arthroplasty for myelopathy versus radiculopathy, And arthroplasty

versus arthrodesis for myelopathy. Clin Spine Surg. 2018 Dec;31(10):420-7. doi: 10.1097/BSD.0000000000000744. PMID: 30371602.

72. Gupta VK, Basantani N, Carvalho AS, et al. Long-term clinicoradiological outcomes of cervical fusion with polyether ether ketone versus cervical disc arthroplasty in a double-blinded randomized control trial. Asian J Neurosurg. 2021 Oct-Dec;16(4):725-31. doi: 10.4103/ajns.AJNS_345_20. PMID: 35071069.

73. Heller JG, Sasso RC, Papadopoulos SM, et al. Comparison of BRYAN cervical disc arthroplasty with anterior cervical decompression and fusion: clinical and radiographic results of a randomized, controlled, clinical trial. Spine. 2009 Jan 15;34(2):101-7. doi: 10.1097/BRS.0b013e31818ee263. PMID: 19112337.

74. Hisey MS, Bae HW, Davis R, et al. Multi-center, prospective, randomized, controlled investigational device exemption clinical trial comparing Mobi-C Cervical Artificial Disc to anterior discectomy and fusion in the treatment of symptomatic degenerative disc disease in the cervical spine. Int J Spine Surg. 2014;8:7. doi: 10.14444/1007. PMID: 25694918.

75. Hisey MS, Bae HW, Davis RJ, et al. Prospective, randomized comparison of cervical total disk replacement versus anterior cervical fusion: results at 48 months follow-up. J Spinal Disord Tech. 2015 May;28(4):E237-43. doi: 10.1097/BSD.0000000000000185. PMID: 25310394.

76. Hisey MS, Zigler JE, Jackson R, et al. Prospective, randomized comparison of one-level mobi-c cervical total disc replacement vs. anterior cervical discectomy and fusion: Results at 5-year follow-up. Int J Spine Surg. 2016;10:10. doi: 10.14444/3010. PMID: 27162712.

77. Hou Y, Nie L, Pan X, et al. Effectiveness and safety of Mobi-C for treatment of single-level cervical disc spondylosis: a randomised control trial with a minimum of five years of follow-up. Bone Joint J. 2016 Jun;98-B(6):829-33. doi: 10.1302/0301-620X.98B6.36381. PMID: 27235528.

78. Jackson RJ, Davis RJ, Hoffman GA, et al. Subsequent surgery rates after cervical total disc replacement using a Mobi-C Cervical Disc Prosthesis versus anterior cervical discectomy and fusion: a prospective randomized clinical trial with 5-year follow-up. J Neurosurg Spine. 2016 May;24(5):734-45. doi: 10.3171/2015.8.SPINE15219. PMID: 26799118.

79. Janssen ME, Zigler JE, Spivak JM, et al. ProDisc-C total disc replacement versus anterior cervical discectomy and fusion for single-level symptomatic cervical disc disease: seven-year follow-up of the prospective randomized u.S. Food and drug administration investigational device exemption study. J Bone Joint Surg Am. 2015 Nov 04;97(21):1738-47. doi: 10.2106/JBJS.N.01186. PMID: 26537161.

80. Karabag H, Cakmak E, Celik B, et al. Arthroplasty versus fusion for single-level cervical disc disease. JPMA J Pak Med Assoc. 2014 Dec;64(12):1348-51. PMID: 25842575.

81. Lanman TH, Burkus JK, Dryer RG, et al. Long-term clinical and radiographic outcomes of the Prestige LP artificial cervical disc replacement at 2 levels: results from a prospective randomized controlled clinical trial. J Neurosurg Spine. 2017 Jul;27(1):7-19. doi: 10.3171/2016.11.SPINE16746. PMID: 28387616.

82. Lavelle WF, Riew KD, Levi AD, et al. Ten-year outcomes of cervical disc replacement with the bryan cervical disc: results from a prospective, randomized, controlled clinical trial. Spine. 2019 May 01;44(9):601-8. doi: 10.1097/BRS.0000000000002907. PMID: 30325888.

83. Loidolt T, Kurra S, Riew KD, et al. Comparison of adverse events between cervical disc arthroplasty and anterior cervical discectomy and fusion: a 10-year follow-up. Spine J. 2021 02;21(2):253-64. doi: 10.1016/j.spinee.2020.10.013. PMID: 33080376.

84. Mummaneni PV, Burkus JK, Haid RW, et al. Clinical and radiographic analysis of cervical disc arthroplasty compared with allograft fusion: a randomized controlled clinical trial. J Neurosurg Spine. 2007 Mar;6(3):198-209. doi: 10.3171/spi.2007.6.3.198. PMID: 17355018.

85. Murrey D, Janssen M, Delamarter R, et al. Results of the prospective, randomized, controlled multicenter Food and Drug Administration investigational device exemption study of the ProDisc-C total disc replacement versus anterior discectomy and fusion for the treatment of 1-level symptomatic cervical disc disease. Spine J. 2009 Apr;9(4):275-86. doi: 10.1016/j.spinee.2008.05.006. PMID: 18774751.

86. Nabhan A, Steudel WI, Nabhan A, et al. Segmental kinematics and adjacent level degeneration following disc replacement versus fusion: RCT with three years of follow-up. J Long Term Eff Med Implants. 2007;17(3):229-36. doi: 10.1615/jlongtermeffmedimplants.v17.i3.60. PMID: 19023947.

87. Nabhan A, Ahlhelm F, Shariat K, et al. The ProDisc-C prosthesis: clinical and radiological experience 1 year after surgery. Spine. 2007 Aug 15;32(18):1935-41. doi: 10.1097/BRS.0b013e31813162d8. PMID: 17700437.

88. Nabhan A, Ishak B, Steudel WI, et al. Assessment of adjacent-segment mobility after cervical disc replacement versus fusion: RCT with 1 year's results. Eur Spine J. 2011 Jun;20(6):934-41. doi: 10.1007/s00586-010-1588-2. PMID: 21221666.

89. Peng-Fei S, Yu-Hua J. Cervical disc prosthesis replacement and interbody fusion: a comparative study. Int Orthop. 2008 Feb;32(1):103-6. doi: 10.1007/s00264-006-0287-4. PMID: 17180356.

90. Phillips FM, Geisler FH, Gilder KM, et al. Long-term outcomes of the US FDA IDE prospective, randomized

controlled clinical trial comparing PCM cervical disc arthroplasty with anterior cervical discectomy and fusion. Spine. 2015 May 15;40(10):674-83. doi: 10.1097/BRS.0000000000000869. PMID: 25955086.

91. Phillips FM, Lee JY, Geisler FH, et al. A prospective, randomized, controlled clinical investigation comparing PCM cervical disc arthroplasty with anterior cervical discectomy and fusion. 2-year results from the US FDA IDE clinical trial. Spine. 2013 Jul 01;38(15):E907-18. doi: 10.1097/BRS.0b013e318296232f. PMID: 23591659.

92. Radcliff K, Coric D, Albert T. Five-year clinical results of cervical total disc replacement compared with anterior discectomy and fusion for treatment of 2-level symptomatic degenerative disc disease: a prospective, randomized, controlled, multicenter investigational device exemption clinical trial. J Neurosurg Spine. 2016 Aug;25(2):213-24. doi: 10.3171/2015.12.SPINE15824. PMID: 27015130.

93. Radcliff K, Davis RJ, Hisey MS, et al. Long-term evaluation of cervical disc arthroplasty with the Mobi-C cervical disc: a randomized, prospective, multicenter clinical trial with seven-year follow-up. Int J Spine Surg. 2017 Netherlands ISASS (E-mail: info@ISASS;11(4):244-62. doi: 10.14444/4031.

94. Sasso RC, Anderson PA, Riew KD, et al. Results of cervical arthroplasty compared with anterior discectomy and fusion: four-year clinical outcomes in a prospective, randomized controlled trial. J Bone Joint Surg Am. 2011 Sep 21;93(18):1684-92. doi: 10.2106/JBJS.J.00476. PMID: 21938372.

95. Vaccaro A, Beutler W, Peppelman W, et al. Long-term clinical experience with selectively constrained SECURE-C cervical artificial disc for 1-level cervical disc disease: results from seven-year follow-up of a prospective, randomized, controlled investigational device exemption clinical trial. Int J Spine Surg. 2018 Jun;12(3):377-87. doi: 10.14444/5044. PMID: 30276095.

96. Vaccaro A, Beutler W, Peppelman W, et al. Clinical outcomes with selectively constrained SECURE-C cervical disc arthroplasty: two-year results from a prospective, randomized, controlled, multicenter investigational device exemption study. Spine. 2013 Dec 15;38(26):2227-39. doi: 10.1097/BRS.0000000000000031. PMID: 24335629.

97. Yang W, Si M, Hou Y, et al. Superiority of 2-level total disk replacement using a cervical disk prosthesis versus anterior cervical diskectomy and fusion. Orthopedics. 2018 Nov 01;41(6):344-50. doi: 10.3928/01477447-20180815-01. PMID: 30125034.

98. Zhang X, Zhang X, Chen C, et al. Randomized, controlled, multicenter, clinical trial comparing BRYAN cervical disc arthroplasty with anterior cervical decompression and fusion in China. Spine (Phila Pa 1976). 2012 Mar 15;37(6):433-8. doi: 10.1097/BRS.0b013e31822699fa. PMID: 21673620.

99. Zhang HX, Shao YD, Chen Y, et al. A prospective, randomised, controlled multicentre study comparing cervical disc replacement with anterior cervical decompression and fusion. Int Orthop. 2014 Dec;38(12):2533-41. doi: 10.1007/s00264-014-2497-5. PMID: 25209344.

100. Zigler JE, Delamarter R, Murrey D, et al. ProDisc-C and anterior cervical discectomy and fusion as surgical treatment for single-level cervical symptomatic degenerative disc disease: five-year results of a Food and Drug Administration study. Spine. 2013 Feb 01;38(3):203-9. doi: 10.1097/BRS.0b013e318278eb38. PMID: 23080427.

101. Donk RD, Arnts H, Verhagen WIM, et al. Cervical sagittal alignment after different anterior discectomy procedures for single-level cervical degenerative disc disease: randomized controlled trial. Acta Neurochir (Wien). 2017 12;159(12):2359-65. doi: 10.1007/s00701-017-3312-z. PMID: 28887690.

102. Coric D, Guyer RD, Bae H, et al. Prospective, multicenter study of 2-level cervical arthroplasty with a PEEK-on-ceramic artificial disc. J Neurosurg Spine. 2022 Apr 01;37(3):1-11. doi: 10.3171/2022.1.SPINE211264. PMID: 35364570.

103. Phillips FM, Coric D, Sasso R, et al. Prospective, multicenter clinical trial comparing M6-C compressible six degrees of freedom cervical disc with anterior cervical discectomy and fusion for the treatment of single-level degenerative cervical radiculopathy: 2-year results of an FDA investigational device exemption study. Spine J. 2020 02;21(2):239-52. doi: 10.1016/j.spinee.2020.10.014. PMID: 33096243.

104. Ostrov PB, Reddy AK, Ryoo JS, et al. Anterior cervical discectomy and fusion versus cervical disc arthroplasty: a comparison of national trends and outcomes. World Neurosurg. 2022 Apr;160:e96-e110. doi: 10.1016/j.wneu.2021.12.099. PMID: 34973439.

105. Nandyala SV, Marquez-Lara A, Fineberg SJ, et al. Comparison between cervical total disc replacement and anterior cervical discectomy and fusion of 1 to 2 levels from 2002 to 2009. Spine. 2014 Jan 01;39(1):53-7. doi: 10.1097/BRS.0000000000000044. PMID: 24108292.

106. Kelly MP, Eliasberg CD, Riley MS, et al. Reoperation and complications after anterior cervical discectomy and fusion and cervical disc arthroplasty: a study of 52,395 cases. Eur Spine J. 2018 06;27(6):1432-9. doi: 10.1007/s00586-018-5570-8. PMID: 29605899.

107. Bhashyam N, De la Garza Ramos R, Nakhla J, et al. Thirty-day readmission and reoperation rates after single-level anterior cervical discectomy and fusion versus those after cervical disc replacement. Neurosurg Focus. 2017 Feb;42(2):E6. doi: 10.3171/2016.11.Focus16407. PMID: 28142261.

108. Grob D, Porchet F, Kleinstuck FS, et al. A comparison of outcomes of cervical disc arthroplasty and fusion in everyday clinical practice: surgical

and methodological aspects. Eur Spine J. 2010 Feb;19(2):297-306. doi: 10.1007/s00586-009-1194-3. PMID: 19882177.

109. Nunley PD, Kerr EJ, 3rd, Cavanaugh DA, et al. Adjacent segment pathology after treatment with cervical disc arthroplasty or anterior cervical discectomy and fusion, part 2: clinical results at 7-year follow-up. Int J Spine Surg. 2020 Jun;14(3):278-85. doi: 10.14444/7037. PMID: 32699748.

110. U.S. Food and Drug Administration. Summary of Safety and Effectiveness Data (SSED). Mobi-C® Cervical Disc Prosthesis (One-level Indication). PMA No.: P110002. August 7, 2013 https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P110002.

111. U.S. Food and Drug Administration. Summary of Safety and Effectiveness Data (SSED). BRYAN® Cervical Disc Prosthesis. PMA No.: P060023. May 12, 2009 https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P060023.

112. U.S. Food and Drug Administration. Summary of Safety and Effectiveness Data (SSED). NuVasive PCM® Cervical Disc System. PMA No.: P100012. October 26, 2012. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P100012.

113. U.S. Food and Drug Administration. Summary of Safety and Effectiveness Data (SSED). PRESTIGE® Cervical Disc System. PMA No.: P060018. July 16, 2007. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P060018.

114. U.S. Food and Drug Administration. Summary of Safety and Effectiveness Data (SSED). ProDisc™-C Total Disc Replacement. PMA No.: P070001. December 17, 2007. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P070001.

115. U.S. Food and Drug Administration. Summary of Safety and Effectiveness Data (SSED). SECURE®-C Artificial Cervical Disc. PMA No.: P100003. September 29, 2012. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P100003.

116. U.S. Food and Drug Administration. Summary of Safety and Effectiveness Data (SSED). M6-C™ Artificial Cervical Disc. PMA No.: P170036. February 6, 2019. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P170036.

117. U.S. Food and Drug Administration. Summary of Safety and Effectiveness Data (SSED). Mobi-C® Cervical Disc Prosthesis (Two-level Indication). PMA No.: P110009. August,23, 2013 https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P110009.

118. U.S. Food and Drug Administration. Summary of Safety and Effectiveness Data (SSED). PRESTIGE LP™ Cervical Disc. PMA No.: P090029. July 7, 2016. https://www.accessdata.fda.gov/scrip

ts/cdrh/cfdocs/cfpma/pma.cfm?id=P090029.

119. U.S. Food and Drug Administration. Summary of Safety and Effectiveness Data (SSED). Simplify® Cervical Artificial Disc. PMA No.: P200022. September 18, 2020. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P200022.

120. Zhou J, Li J, Lin H, et al. Could self-locking stand-alone cage reduce adjacent-level ossification development after aneterior cervical discectomy and fusion? J Clin Neurosci. 2020 Aug;78:60-6. doi: 10.1016/j.jocn.2020.06.014. PMID: 32624365.

121. Scholz M, Onal B, Schleicher P, et al. Two-level ACDF with a zero-profile stand-alone spacer compared to conventional plating: a prospective randomized single-center study. Eur Spine J. 2020 11;29(11):2814-22. doi: 10.1007/s00586-020-06454-z. PMID: 32430769.

122. Panchal RR, Kim KD, Eastlack R, et al. A clinical comparison of anterior cervical plates versus stand-alone intervertebral fusion devices for single-level anterior cervical discectomy and fusion procedures. World Neurosurg. 2017 Mar;99:630-7. doi: 10.1016/j.wneu.2016.12.060. PMID: 28017756.

123. Chen Y, Chen H, Wu X, et al. Comparative analysis of clinical outcomes between zero-profile implant and cages with plate fixation in treating multilevel cervical spondilotic myelopathy: A three-year follow-up. Clin Neurol Neurosurg. 2016 May;144:72-6. doi: 10.1016/j.clineuro.2016.03.010. PMID: 26999528.

124. Li Y, Hao D, He B, et al. The efficiency of zero-profile implant in anterior cervical discectomy fusion: a prospective controlled long-term follow-up study. J Spinal Disord Tech. 2015 Dec;28(10):398-403. doi: 10.1097/BSD.0000000000000032. PMID: 24136051.

125. He S, Feng H, Lan Z, et al. A randomized trial comparing clinical outcomes between zero-profile and traditional multilevel anterior cervical discectomy and fusion surgery for cervical myelopathy. Spine. 2018 03 01;43(5):E259-E66. doi: 10.1097/BRS.0000000000002323. PMID: 29432408.

126. Nemoto O, Kitada A, Naitou S, et al. Stand-alone anchored cage versus cage with plating for single-level anterior cervical discectomy and fusion: a prospective, randomized, controlled study with a 2-year follow-up. Eur. 2015 Jul;25 Suppl 1:S127-34. doi: 10.1007/s00590-014-1547-4. PMID: 25283362.

127. Zhang B, Jiang YZ, Song QP, et al. Outcomes of cervical degenerative disc disease treated by anterior cervical discectomy and fusion with self-locking fusion cage. World j. 2022;10(15):4776-84. doi: 10.12998/wjcc.v10.i15.4776. PMID: 35801046.

128. Zavras AG, Nolte MT, Sayari AJ, et al. Stand-alone cage versus anterior plating for 1-level and 2-level anterior cervical discectomy and fusion: a randomized controlled trial. Clin Spine Surg. 2022 05

01;35(4):155-65. doi: 10.1097/BSD.0000000000001332. PMID: 35394961.

129. Chen Y, Wang X, Lu X, et al. Comparison of titanium and polyetheretherketone (PEEK) cages in the surgical treatment of multilevel cervical spondylotic myelopathy: a prospective, randomized, control study with over 7-year follow-up. Eur Spine J. 2013 Jul;22(7):1539-46. doi: 10.1007/s00586-013-2772-y. PMID: 23568254.

130. Niu CC, Liao JC, Chen WJ, et al. Outcomes of interbody fusion cages used in 1 and 2-levels anterior cervical discectomy and fusion: titanium cages versus polyetheretherketone (PEEK) cages. J Spinal Disord Tech. 2010 Jul;23(5):310-6. doi: 10.1097/BSD.0b013e3181af3a84. PMID: 20124907.

131. Godlewski B, Bebenek A, Dominiak M, et al. PEEK versus titanium-coated PEEK cervical cages: fusion rate. Acta Neurochir (Wien). 2022 06;164(6):1501-7. doi: 10.1007/s00701-022-05217-7. PMID: 35471708.

132. Kanna RM, Perambuduri AS, Shetty AP, et al. A randomized control trial comparing local autografts and allografts in single level anterior cervical discectomy and fusion using a stand-alone cage. Asian spine j. 2021 Dec;15(6):817-24. doi: 10.31616/asj.2020.0182. PMID: 33189111.

133. Arnold PM, Sasso RC, Janssen ME, et al. i-Factor TM bone graft vs autograft in anterior cervical discectomy and fusion: 2-year follow-up of the randomized single-blinded food and drug administration investigational device exemption study. Neurosurgery; 2018. p. 377-84.

134. Xie Y, Li H, Yuan J, et al. A prospective randomized comparison of PEEK cage containing calcium sulphate or demineralized bone matrix with autograft in anterior cervical interbody fusion. Int Orthop. 2015 Jun;39(6):1129-36. doi: 10.1007/s00264-014-2610-9. PMID: 25432324.

135. Cho DY, Lee WY, Sheu PC, et al. Cage containing a biphasic calcium phosphate ceramic (Triosite) for the treatment of cervical spondylosis. Surg Neurol. 2005 Jun;63(6):497-503; discussion -4. doi: 10.1016/j.surneu.2004.10.016. PMID: 15936361.

136. Baskin DS, Ryan P, Sonntag V, et al. A prospective, randomized, controlled cervical fusion study using recombinant human bone morphogenetic protein-2 with the CORNERSTONE-SR allograft ring and the ATLANTIS anterior cervical plate. Spine. 2003 Jun 15;28(12):1219-24; discussion 25. doi: 10.1097/01.BRS.0000065486.22141.CA. PMID: 12811263.

137. Yi J, Lee GW, Nam WD, et al. A prospective randomized clinical trial comparing bone union rate following anterior cervical discectomy and fusion using a polyetheretherketone cage: hydroxyapatite/b-tricalcium phosphate mixture versus hydroxyapatite/demineralized bone matrix mixture. Asian spine j. 2015 Feb;9(1):30-8. doi:

10.4184/asj.2015.9.1.30. PMID: 25705332.

138. Arnold PM, Anderson KK, Selim A, et al. Heterotopic ossification following single-level anterior cervical discectomy and fusion: results from the prospective, multicenter, historically controlled trial comparing allograft to an optimized dose of rhBMP-2. J Neurosurg Spine. 2016 Sep;25(3):292-302. doi: 10.3171/2016.1.SPINE15798. PMID: 27129045.

139. Smucker JD, Rhee JM, Singh K, et al. Increased swelling complications associated with off-label usage of rhBMP-2 in the anterior cervical spine. Spine (Phila Pa 1976). 2006 Nov 15;31(24):2813-9. doi: 10.1097/01.brs.0000245863.52371.c2. PMID: 17108835.

140. Vedantam A, Rajshekhar V. Does the type of T2-weighted hyperintensity influence surgical outcome in patients with cervical spondylotic myelopathy? A review. Eur Spine J. 2013 Jan;22(1):96-106. doi: 10.1007/s00586-012-2483-9. PMID: 22926434.

141. Aggarwal RA, Srivastava SK, Bhosale SK, et al. Prediction of surgical outcome in compressive cervical myelopathy: a novel clinicoradiological prognostic score. J Craniovertebr Junction Spine. 2016 Apr-Jun;7(2):82-6. doi: 10.4103/0974-8237.181828. PMID: 27217653.

142. Baker JD, Harada GK, Tao Y, et al. The impact of modic changes on preoperative symptoms and clinical outcomes in anterior cervical discectomy and fusion patients. Neurospine. 2020 Mar;17(1):190-203. doi: 10.14245/ns.2040062.031. PMID: 32252168.

143. Harada GK, Alter K, Nguyen AQ, et al. Cervical spine endplate abnormalities and association with pain, disability, and adjacent segment degeneration after anterior cervical discectomy and fusion. Spine. 2020 Aug 01;45(15):E917-E26. doi: 10.1097/BRS.0000000000003460. PMID: 32675603.

144. Kim TH, Ha Y, Shin JJ, et al. Signal intensity ratio on magnetic resonance imaging as a prognostic factor in patients with cervical compressive myelopathy. Medicine (Baltimore). 2016 Sep;95(39):e4649. doi: 10.1097/MD.0000000000004649. PMID: 27684796.

145. Li XY, Lu SB, Sun XY, et al. Clinical and magnetic resonance imaging predictors of the surgical outcomes of patients with cervical spondylotic myelopathy. Clin Neurol Neurosurg. 2018 11;174:137-43. doi: 10.1016/j.clineuro.2018.09.003. PMID: 30241007.

146. Nouri A, Martin AR, Kato S, et al. The relationship between MRI signal intensity changes, clinical presentation, and surgical outcome in degenerative cervical myelopathy: Analysis of a global cohort. Spine. 2017 Dec 15;42(24):1851-8. doi: 10.1097/BRS.0000000000002234. PMID: 28498290.

147. Sarkar S, Turel MK, Jacob KS, et al. The evolution of T2-weighted intramedullary signal changes following ventral decompressive surgery for cervical spondylotic myelopathy: Clinical article. J

Neurosurg Spine. 2014 Oct;21(4):538-46. doi: 10.3171/2014.6.SPINE13727. PMID: 25014501.

148. Sharma R, Borkar S, Katiyar V, et al. Interplay of dynamic extension reserve and T1 slope in determining the loss of cervical lordosis following laminoplasty: a novel classification system. World Neurosurg. 2020 Apr;136:e33-e40. doi: 10.1016/j.wneu.2019.08.212. PMID: 31493608.

149. Uchida K, Nakajima H, Takeura N, et al. Prognostic value of changes in spinal cord signal intensity on magnetic resonance imaging in patients with cervical compressive myelopathy. Spine J. 2014 Aug 01;14(8):1601-10. doi: 10.1016/j.spinee.2013.09.038. PMID: 24411833.

150. Yin LQ, Zhang J, Wu YG, et al. Increased signal intensity of spinal cord on T2W magnetic resonance imaging for cervical spondylotic myelopathy patients: Risk factors and prognosis (a STROBE-compliant article). Medicine (Baltimore). 2020 Dec 04;99(49):e23098. doi: 10.1097/MD.0000000000023098. PMID: 33285685.

151. Zhang JT, Meng FT, Wang S, et al. Predictors of surgical outcome in cervical spondylotic myelopathy: focusing on the quantitative signal intensity. Eur Spine J. 2015 Dec;24(12):2941-5. doi: 10.1007/s00586-015-4109-5. PMID: 26155898.

152. Zhang P, Shen Y, Zhang YZ, et al. Significance of increased signal intensity on MRI in prognosis after surgical intervention for cervical spondylotic myelopathy. J Clin Neurosci. 2011 Aug;18(8):1080-3. doi: 10.1016/j.jocn.2010.12.023. PMID: 21696960.

153. Morio Y, Teshima R, Nagashima H, et al. Correlation between operative outcomes of cervical compression myelopathy and MRI of the spinal cord. Spine. 2001 Jun 01;26(11):1238-45. PMID: 11389390.

154. Wang K, Chen Z, Zhang F, et al. Evaluation of DTI parameter ratios and diffusion tensor tractography grading in the diagnosis and prognosis prediction of cervical spondylotic myelopathy. Spine (Phila Pa 1976). 2017 Feb 15;42(4):E202-e10. doi: 10.1097/brs.0000000000001784. PMID: 28207659.

155. Suri A, Chabbra RP, Mehta VS, et al. Effect of intramedullary signal changes on the surgical outcome of patients with cervical spondylotic myelopathy. Spine J. 2003 Jan-Feb;3(1):33-45. doi: 10.1016/s1529-9430(02)00448-5. PMID: 14589243.

156. Fukushima T, Ikata T, Taoka Y, et al. Magnetic resonance imaging study on spinal cord plasticity in patients with cervical compression myelopathy. Spine (Phila Pa 1976). 1991 Oct;16(10 Suppl):S534-8. doi: 10.1097/00007632-199110001-00016. PMID: 1801267.

157. Lambrechts MJ, D'Antonio ND, Karamian BA, et al. What is the role of dynamic cervical spine radiographs in predicting pseudarthrosis revision following anterior cervical discectomy and fusion? Spine J. 2022 May

12;12(10):12. doi: 10.1016/j.spinee.2022.04.020. PMID: 35568109.

158. Song KS, Piyaskulkaew C, Chuntarapas T, et al. Dynamic radiographic criteria for detecting pseudarthrosis following anterior cervical arthrodesis. J Bone Joint Surg Am. 2014 Apr 02;96(7):557-63. doi: 10.2106/JBJS.M.00167. PMID: 24695922.

159. Badhiwala JH, Nassiri F, Witiw CD, et al. Investigating the utility of intraoperative neurophysiological monitoring for anterior cervical discectomy and fusion: analysis of over 140,000 cases from the National (Nationwide) Inpatient Sample data set. J Neurosurg Spine. 2019 03 29;31(1):76-86. doi: 10.3171/2019.1.SPINE181110. PMID: 30925481.

160. Ajiboye RM, D'Oro A, Ashana AO, et al. Routine use of intraoperative neuromonitoring during ACDFs for the treatment of spondylotic myelopathy and radiculopathy is questionable: a review of 15,395 cases. Spine. 2017 Jan 01;42(1):14-9. doi: 10.1097/BRS.0000000000001662. PMID: 27120059.

161. Smith SS, Stewart ME, Davies BM, et al. The prevalence of asymptomatic and symptomatic spinal cord compression on magnetic resonance imaging: a systematic review and meta-analysis. Global spine j. 2021;11(4):597-607. doi: 10.1177/2192568220934496. PMID: 32677521.

162. Nakashima H, Yukawa Y, Suda K, et al. Narrow cervical canal in 1211 asymptomatic healthy subjects: the relationship with spinal cord compression on MRI. Eur Spine J. 2016 Jul;25(7):2149-54. doi: 10.1007/s00586-016-4608-z.

163. Kovalova I, Kerkovsky M, Kadanka Z, et al. Prevalence and imaging characteristics of nonmyelopathic and myelopathic spondylotic cervical cord compression. Spine (Phila Pa 1976). 2016 Dec 15;41(24):1908-16. doi: 10.1097/brs.0000000000001842.

164. Nouri A, Tessitore E, Molliqaj G, et al. Degenerative Cervical Myelopathy: Development and Natural History [AO Spine RECODE-DCM Research Priority Number 2]. Global spine j. 2022;12(1):39S-54S. doi: 10.1177/21925682211036071.

165. Martin AR, De Leener B, Cohen-Adad J, et al. Can microstructural MRI detect subclinical tissue injury in subjects with asymptomatic cervical spinal cord compression? A prospective cohort study. BMJ Open. 2018 Apr 13;8(4):e019809. doi: 10.1136/bmjopen-2017-019809. PMID: 29654015.

166. Bednarik J, Kadanka Z, Dusek L, et al. Presymptomatic spondylotic cervical myelopathy: an updated predictive model. Eur Spine J. 2008 Mar;17(3):421-31. doi: 10.1007/s00586-008-0585-1. PMID: 18193301.

167. Kadanka Z, Jr., Adamova B, Kerkovsky M, et al. Predictors of symptomatic myelopathy in degenerative cervical spinal cord compression. Brain Behav. 2017 Sep;7(9):e00797. doi: 10.1002/brb3.797. PMID: 28948090.

168. Nurick S. The pathogenesis of the spinal cord disorder associated with cervical spondylosis. Brain. 1972;95(1):87-100. doi: 10.1093/brain/95.1.87. PMID: 5023093.

169. Beattie MS, Manley GT. Tight squeeze, slow burn: inflammation and the aetiology of cervical myelopathy. Brain. 2011 May;134(Pt 5):1259-61. doi: 10.1093/brain/awr088. PMID: 21596766.

170. Karadimas SK, Moon ES, Yu WR, et al. A novel experimental model of cervical spondylotic myelopathy (CSM) to facilitate translational research. Neurobiol Dis. 2013 Jun;54:43-58. doi: 10.1016/j.nbd.2013.02.013. PMID: 23466695.

171. Yu WR, Baptiste DC, Liu T, et al. Molecular mechanisms of spinal cord dysfunction and cell death in the spinal hyperostotic mouse: implications for the pathophysiology of human cervical spondylotic myelopathy. Neurobiol Dis. 2009 Feb;33(2):149-63. doi: 10.1016/j.nbd.2008.09.024. PMID: 19006686.

172. Yu WR, Liu T, Kiehl TR, et al. Human neuropathological and animal model evidence supporting a role for Fas-mediated apoptosis and inflammation in cervical spondylotic myelopathy. Brain. 2011 May;134(Pt 5):1277-92. doi: 10.1093/brain/awr054. PMID: 21490053.

173. Karadimas SK, Erwin WM, Ely CG, et al. Pathophysiology and natural history of cervical spondylotic myelopathy. Spine (Phila Pa 1976). 2013 Oct 15;38(22 Suppl 1):S21-36. doi: 10.1097/BRS.0b013e3182a7f2c3. PMID: 23963004.

174. Martin AR, Jentzsch T, Wilson JRF, et al. Inter-rater reliability of the Modified Japanese Orthopedic ssociation score in degenerative cervical myelopathy: a cross-sectional study. Spine. 2021 Aug 15;46(16):1063-9. doi: 10.1097/BRS.0000000000003956. PMID: 33492085.

175. Yonenobu K, Abumi K, Nagata K, et al. Interobserver and intraobserver reliability of the japanese orthopaedic association scoring system for evaluation of cervical compression myelopathy. Spine (Phila Pa 1976). 2001 Sep 1;26(17):1890-4; discussion 5. doi: 10.1097/00007632-200109010-00014. PMID: 11568701.

176. Fehlings MG, Tetreault LA, Riew KD, et al. A clinical practice guideline for the management of patients with degenerative cervical myelopathy: recommendations for patients with mild, moderate, and severe disease and nonmyelopathic patients with evidence of cord compression. Global spine j. 2017 Sep;7(3 Suppl):70S-83S. doi: 10.1177/2192568217701914. PMID: 29164035.

177. Fehlings MG, Wilson JR, Kopjar B, et al. Efficacy and safety of surgical decompression in patients with cervical spondylotic myelopathy: results of the AOSpine North America prospective multi-center study. J Bone Joint Surg Am. 2013 Sep 18;95(18):1651-8. doi: 10.2106/JBJS.L.00589. PMID: 24048552.

178. Shimomura T, Sumi M, Nishida K, et al. Prognostic factors for deterioration of patients with cervical spondylotic myelopathy after nonsurgical treatment. Spine (Phila Pa 1976). 2007 Oct 15;32(22):2474-9. doi: 10.1097/BRS.0b013e3181573aee. PMID: 18090088.

179. Vonck CE, Tanenbaum JE, Smith GA, et al. National Trends in Demographics and Outcomes Following Cervical Fusion for Cervical Spondylotic Myelopathy. Global spine j. 2018 May;8(3):244-53. doi: 10.1177/2192568217722562. PMID: 29796372.

Abbreviations and Acronyms

ACDF	Anterior Cervical Discectomy and Fusion
ADL	Activities of Daily Living
AHRQ	Agency for Healthcare and Research Quality
AUC	Area Under the Curve
BMP-2	Bone Morphogenetic protein 2
C-ADR	Cervical Artificial Disc Replacement
CCI	Charlson Comorbidity Index
CDD	Cervical Degenerative Disease
CER	Comparative Effectiveness Review
CI	Confidence Interval
CNS	Congress of Neurological Surgeons
DRI	Disability Rating Index
EQ-5Dm	EuroQol-5 dimension instrument
IDE	Investigational Device Exemption
IONM	Intraoperative neuromonitoring
ISI	Increased Signal Intensity
JOA	Japanese Orthopaedic Association Scale
KQ	Key Question
MD	Mean Difference
MDI	Myelopathy Disability Index
mJOA	Modified Japanese Orthopaedic Association Scale
MRI	Magnetic Resonance Imaging
NDI	Neck Disability Index
NIS	National Inpatient Sample
NRS	Numeric Rating Scale
NRSI	Nonrandomized studies of interventions
OPLL	Ossification of the Posterior Longitudinal Ligament
PCDF	Posterior Cervical Decompression and Fusion
PEEK	Polyetheretherketone
PEMF	Pulsed Electro-Magnetic Field
PICOTS	Population, Intervention, Comparator, Outcome, Time, Setting
PROMIS-29	Patient reported outcome measurement information system
RCT	Randomized Controlled Trials
RR	Risk Ratio
SCC	Spinal Cord Compression
SF-12	12-Item Short Form Health Survey

SF-36	36-Item Short Form Health Survey
SIR	Signal Intensity Ratio
SWAL-QOL	Swallowing Quality of Life Questionnaire
VAS	Visual Analog Scale